SearcherJan, 1998

Cited references on the Web: a review of ISI's Web of Science. (Institute for Scientific Information; includes related articles)

Author/s: Deborah Lynne Wiley

The Web of Science is the somewhat inappropriate name the Institute for Scientific Information (ISI) has given to the new Web interface to its citation indexes. While a very nice interface to the ISI databases, megabases that certainly do cover a broad range of scientific disciplines, the "Web" aspect only refers to the fact that you access the data via the World Wide Web, not that it provides a comprehensive Web of information. At least not yet. And, with the Arts & Humanities and Social Science Citation indexès included, the product actually covers more than just science. Ah, but then what's in a name?

Introduced in April 1997, ISI already has dozens of sites accessing their data through the Web. Subscribers have two access choices. They can load the data on their own server in-house and control access from there. This choice often works best for corporations that want data access secured behind their own firewalls. But this choice means user institutions must handle a lot of data; the Science Citation Index Expanded alone runs to more than 5.5 gigabytes per year. For those unwilling or unable to maintain many gigabytes of data and multitudes of users, you can also access the data through the server housed at ISI in Philadelphia. This adds a small technology fee to your subscription, but may be worth it for the convenience of having someone else maintain the server. Access is controlled by use of ranges of IP addresses; remote users get in through proxy servers. ISI plans to set up additional servers in Asia and Europe to help ease access in other parts of the world.

On the client side, you need a standard Web browser. ISI recommends Netscape Navigator version 3.0 or higher, but also tries to support Microsoft Internet Explorer and some older browsers. Certain features simply will not work with particular browsers. The ISI help desk tries to keep track of all the bugs or problems with

each variety of browser. I first tried to access with Netscape Navigator Gold and discovered there a bug in versions 3.02-3.04 that somehow made the server time-out before I could even get in. Navigator 3.0 worked fine.

Content

Web of Science carries the past 10 years of ISI's three major citation indexes: Arts & Humanities Citation Index (A&HCI), Social Science Citation Index (SSCI), and the Science Citation Index Expanded (SCIE). The SCIE covers about 2,000 more journals than the SciSearch database, as it includes all the science journals collected by IST for its various products. The other indexes are the equivalent to current dial-up online files in their content. Combined, the three indexes provide access to article citations from about 8,000 journals and over 10 million records. Each index updates weekly. ISI is currently testing files back to the mid-'70s and, believe it or not, they hope to put up all of the years of the citation indexes, from day one. That's back to 1945 for SCI and 1956 for SSCI -- potentially more complete coverage than any other current service! All the indexes carry cited references or footnotes. What a wealth of information! But will any library really pay year by year back to 1945?
All of the fields available online now also operate on the Web of Science version. Access to some data works better on Web of Science than current online dial-up systems. To accomplish this, ISI has developed their own search engine from scratch -- throwing out attempts to alter a third party engine. At last! An implementation that does full justice to the cited references.

Other Internet Efforts from ISI

In the past few years, ISI has embraced the Internet. They use it as a delivery mechanism for as many of their products as possible, not to mention press releases and product information. Current Contents is available on the Web as Current Contents Connect. The ISI Alerting Services have been expanded to provide e-mail delivery of results from customized profiles searches or a journal tracking service. ISI hopes to extend this onto the Web next year with the development of Discovery Agent, a browser-based interface to let users create, manage, and edit their own profiles. In addition, ISI plans to put its chemistry data on the Web using software from Hampden Data Services Ltd to enable searching and displaying of structural data. Look for this later in 1998. ISI's next step should be to develop some new products specifically designed to harness the latest technologies and the new user behavior of the technology generation.

Pricing

Subscribers must purchase data year by year and index by index. Although ISI is reluctant to quote prices, since so many variables

affect the price, a safe ballpark figure would put one year of the SCIE for somewhere in the low fifties -- that's $50,000 or so for one year of this one index in a system that carries 10 years each of three indexes. Factors affecting the price include number of sites, number of simultaneous users, years of data, current subscriptions to other ISI products, and a variety of other conditions. You pay a one-time fee for each back year and an ongoing annual subscription price for 52 weekly updates. And, if you choose to access through ISI's server, a small technology fee is added (certainly less than what it would cost to store the data in-house). Free trials are available on up to five years of data.

At the moment ISI is still thinking subscription-only pricing. What about transactional pricing? Why not let anyone search and see the summary results, and pay for the full record? Come on ISI -- at least entertain the idea.

Searching

The Web interface to the citation indexes has more functionality and easier access than any other version of ISI products. While reviewing this system, I got access to all three citation databases for the past 10 years. What a treat. Searching to my heart's content -- with no meter running. The biggest benefit of the Web of Science is easy end-user access at a high, but fixed fee. That's not to say the system is perfect. What on the Web is? But as a first generation product, this is very good.
There are two search modes -- Quick Search and Full Search. The Quick Search is, as it says, much quicker. But it also provides limited functionality. Designed for novice user, it does not want to confuse with too many options. The options an experienced searcher wants appear in the full search.

Quick Search Mode

When entering this mode, the system first prompts you to select which databases you want to search, SCIIE, SSCI, and/or A&HCI. You can search them all at once if you want. Then you click on one of three buttons: Topic, Person, or Place. At this point you cannot combine a topic search directly with an author or cited reference search, however, you can certainly follow the trail of the cited references from any record you retrieve.

Topic

If you click Topic, you get a box in which to enter your search terms. The default treats multiple words as a phrase. You can also use the Boolean operators AND, OR, NOT, and SAME (in the same sentence), and parse the search query using parentheses (nesting). The results display in groups of 10, showing the title, author, and source. Click on the hyperlinked title to display the "full record" which is not full-text, but the author's abstract, if available. This full record states the

number of references cited at the end of the article and also indicates if this article itself has been cited. Clicking on these links displays the appropriate list of references. Any article listed in the references that also appears referenced in the database is hyperlinked for instant access. This linking of cited references is done through a proprietary algorithm and may prove useful in the future as ISI tries to link to full-text sources. But more about that later.

Person

Clicking the Person button in Quick Search presents a box in which you can enter a person's name. Conveniently, it states right above this box the format that you should use -- last name space initials. Searching on a last name alone retrieves all occurrences of that name. Last name plus one initial plus gives just one initial. Last name plus one initial plus(*) gives truncation to get one or more initials. Radio buttons then let you choose whether you want to search for this person 1) as an author of a paper; 2) as a subject of the paper; or 3) to get all the articles in the database that cite this person's work. What a perfect way to phrase the question for those unfamiliar with cited reference searching. Click search and a list of results appears as above, with all the same functionality.

Place

A place search allows you to seek work being done at a particular institution or geographic region. The search automatically limits to the author address field. You get a box to enter your data, and help screens provide advice and examples for the many abbreviations used in the address field.

Using the Results

Displaying and Marking Records

After a search, up to 10 results at a time display in a summary list on an HTML page, providing title, first three authors, and source. The title hyperlinks to the full record, which includes all the named authors and a link to cited references. While displaying the summary list of records, you can mark individual records for later printing or downloading. Don't forget to click the Submit button. The system can't recognize your marks until you submit the page back to the server. Awkward, but necessary at this point.

There is also a button that says "Mark All" but should say "Mark All Records on This Page." It marks the records displayed on that one page. If the search has retrieved more than 10 results, the user must click Next to see the next page, and then must mark the records on that page. This forces the user to actually look at the results before just saving dozens of irrelevant records. At this point, the system supplies no way to display all the records at one time in full record format, i.e., with abstracts. You must view them one at a

time. Nor can you see a long list of all your results in summary format. There you must go 10 at a time. ISI really should offer more flexibility in designating the number of records you want displayed and in which format. However, for your set of marked records, you can choose the fields you want printed (the default is title, author, and source fields) and then view all your marked records on the screen, formatted for printing.

Printing and Downloading Records

In theory you can print all fields, but in my testing if you have too much data for the server to process in the time allocated, you get nothing but a server effort message. I could get a set of 26 records to print with abstracts, but not with cited references -- with or without abstracts. In all probability a simple correction on the server side could change the timeout limits.

Once the system compiles the records for printing, you can view them all in one long file on your screen, with all your chosen fields. This represents, more or less, what you will get when you print that file. A note about printing. Success can very much depend on your browser. Some of the older versions of Netscape will print each record on a separate page. The newer versions don't do this, but may start a new page if the entire record doesn't fit on the page. Microsoft Internet Explorer may also have problems. Browsers are obviously not designed for optimal printing.

Other choices for output involve the Save to File in tagged format (check the help screens for the complete list of two-letter tags) or an export to Pro-cite or Reference Manager formats. (These two bibliographic management software packages are owned by Research Information Systems, a subsidiary of ISI). If you download you can get all of the fields with no problem, much faster than trying to print.

Full Search Mode

Click Full Search and you get the same opportunity as in Quick Search mode to choose your database. But here you can also choose which years you want to search. The Quick Search automatically searches all the years of your subscription. You then choose between General Search and Cited Reference searching. Unfortunately, in a major limitation for the product at present, you cannot combine these two types of searching. To search for an author and all the papers that cited them takes two steps, although ISI hopes to implement combined searching capabilities sometime in the future. But, you can combine the sets of marked records from each type of search into one file for printing or downloading. In fact, the system retains your marked records throughout all your searches until you start a new session; it does not clear out the marked records after every new search. You can switch between Quick and Full Search modes and do any number of searches, and the system will keep all

your marked records in one file. To clear the marked records you must either manually de-select each record from your list of marked records, or go back to the Web of Science entry page and click New Session.

General Search

This option offers boxes where the searcher can enter words for topic, author, source title, and address searches. A topic search will search all the words in the title, abstract, author's keywords, and ISI's automatically generated KeyWords Plus. KeyWords Plus consists of words or phrases that frequently appear in the titles of an article's references, but do not appear in the title of the article itself. A checkbox lets you limit the search to just the tide. An author search seeks through all the authors of an article, no matter how many.

The source field includes the name of the original source document, which may be heavily abbreviated. There is an option to select from the complete list of journals to get the exact name. This list is long -- over 8,000 journals broken into alphabetic segments, and you must copy and paste (or type) the name from the list into the search box. ISI should really make this feature easier to handle.

The address search looks only in the author address field, where again data appears heavily abbreviated. The Help screens give a long list of common abbreviations to help you find what you want. You can enter data in any or all of the fields, using Boolean operators, truncation, and parentheses to parse the query. You can limit your search to a particular language or document type by selecting your choice from a scroll box. You can also sort the results by latest date (the default), relevance, first author, or source title. You cannot double sort on multiple levels, e.g., an author sort within a source sort. When you've typed or selected all your options, click Search and the summary list of results appears, 10 at a time, exactly the same as in Quick Search.

Cited Reference Searching

Of course, citation searching is what the citation indexes are all about. The Full Search mode lets you specify exactly which citation you want to use. Quick Search simply retrieves all articles that cite a particular author. Click the Cited Reference Search button to get a screen with boxes to specify author, source title, and year. Enter as much data as you want, and click Look Up. For records in the database, you can search by any of the named authors (something not possible in the SciSearch online files where you can only search by first author). For cited references whose corresponding record does not appear in the database, only the first named author may be searched. All the cited references matching your criteria display 10 at a time. This allows you to find variations in the citation or to select multiple citations. You can mark the references you want to use and then click Search. This retrieves all the records in the

database that have cited your selected references. These will display in the standard summary format, 10 at a time.

Related Records

Besides citation searching, the system offers additional ways to get articles that share references. For each article with a list of references attached, you can click on the Related Records link to retrieve other articles that cite at least one of the same references. These sort by relevancy so that articles with the most cited references in common appear at the top of the list. You can use the entire list of cited references at the end of your article to find related records, or you can mark particular references to use. The default is to use all the references, so they are all marked. If you want to use just a few citations from a list of references, you must unmark each of the unwanted references individually. This can be a bit tedious with a long list of references. A button that lets you "unmark all" would help a lot.
The related records searching technique lets you surf laterally through the database. You can go through several levels of related records, just by a click of a button. Many scientists will swear that this is the best way to track trends in science, by following the papers that the experts find important to their work. And you can view the database record of many references just by clicking the hyperlink. The danger is that it is easy to get lost. Follow too many references and you forget where you started. The browser's Back key will take you back a screen at a time, but you still may not easily find where you started.

Cited reference searching also lets you go forward in time. In any full record you retrieve you can see the number of records in the database that have cited that article. Click on the Times Cited number and you get a summary list of those records. From there you can view individual records and their cited references, and so on. If you want, you can spend a lot of time browsing through subjects via related records and cited references. This product begins to fulfill ISI founder Eugene Garfield's vision of what citation analysis makes possible. What an aid to the "serendipity" factor that scientists demand! No other database provides this kind of forward and lateral movement.

Saving Your Query

The Full Search mode provides an opportunity to save your search query. You cannot save queries in Quick Search. Click the Save Query button and the browser Save As box opens. ISI recommends that you save die file with an HTML extension, because browsers default to this file type. You can save the file to any local disk, directory, or folder; if you have your own server, you can store it there. To rerun the search you need to use the File, Open File feature of the browser to open that page. You then see the search page with all your search entries typed in. Unfortunately, nothing on

the screen prompts you to run your saved query, and you must remember the file name and path of where you saved the search. Since the operation occurs through the browser itself, not the system, there is no button to click to recall your search. Perhaps a hyperlink, worded, "Run your saved search," that just links to the help pages which describe how to do it, would help.

Document Delivery

At the moment, no full-text sources link to this product. However, you can link the Web of Science to your Web-based OPAC so users can click into your journal holdings, and presumably into the full text, if it just so happens you have access to the electronic version of a journal. Currently the link to holdings, operates at the journal level, but as publishers' sites create linkages at the article level, this could certainly connect to in-house data as well. Also, you can forward a set of marked records directly to a document delivery service or a gateway. You can specify an e-mail address to receive a listing. Naturally, the system defaults to ISI's Document Solution delivery service. Every record contains the ISI document delivery number. Simply click on the Format for Document Delivery button, enter your name and account number, and your list of articles goes to ISI's Document Solution.
Unfortunately, you get no warning of what the documents will cost. Unlike Ovid or OCLC and other services, Web of Science offers no way to distinguish what a library subscriber owns. ISI has plans underway to work on a journal level approach to this problem, linking to library OPACs, but as yet they do not have plans to provide article-level notifications.

Full Text

ISI knows that they need to provide immediate access to the full text. They are working feverishly to get publishers to link their data. They have announced a linkage from the journal Nature to the Web of Science scheduled for early 1998, but only for subscribers of both products. This will presumably work in a similar fashion to the linkages from the online version of Science magazine to NLM's PubMed database [see sidebars]. In addition, Jacqueline Trolley, Director of Corporate Communications at ISI, has assured me that "by mid-1998 we expect a significant number of linkages to the primary publishing community."

Ultimately, ISI's vision is to automatically generate a unique identifier that links the full-text article with the database record. This identifier would not follow any standards currently under discussion, such as the Digital Object Identifier (DOI). The National Library of Medicine's PubMed does this now with their associated publishers. NLM generates a unique number to code an article and then sends it back to the publisher. The publisher's job is to put their full text on the Web and include the unique identifier with each appropriate article. Again like PubMed, ISI envisions the user's access to the full-

text journal will depend on having a subscription to that journal.

One part of the task involves linking the database record to the article. Ideally you would also want the references at the end of each article linked back though the database to the abstract and then again to the full text of the referenced article. The more difficult task comes with finding a way to keep the user from having to log in to every full-text site from all the different publishers. As more publishers put their journals on the Web. access from a broad, multidisciplinary database will become essential to help the poor user connect to all the different sites. Seems a win-win situation for all concerned, so it is puzzling why primary publishers are so hesitant to sign on. Currently some organizations that subscribe to electronic journals from various sources and to the Web of Science have begun threatening to link the two delivery routes themselves, if ISI and the publishers can't get their act together.

Sometimes I hate the word subscription. As with access to the Web of Science itself, ISI intends only to let a user link to a site where they have a subscription to the full text. How hard would it be to implement transaction pricing on the article level? Or does this threaten the publisher's concept of a journal?

Also, libraries should be aware that the standard subscription agreement with ISI is a lease. You do not have any right to the data should you cancel the subscription. However, for about 20 percent more, you can license rather than lease the data. Then if you cancel, ISI will give you a copy of the data you already paid for and the current version of the software. You still have to have a place to store the data and provide the means of access.

Documentation and Help

The Web of Science provides very thorough and well-written help screens. It links to in-context help in various areas, especially related to examples of search techniques or field peculiarities. You can also pull all the help pages together in a format suitable for printing. In addition, ISI's Education Department will come and do training at your site -- for both librarians and end users. They have developed detailed documentation and may make the material available over the Web. Right now you can access a good primer on cited reference searching from the ISI corporate Web site, but not directly from the Web of Science. Although the Web of Science interface is easy to use, even the most experienced searchers can pick up helpful tips from ISI's experts.

What's Next?

I like this product. When I was in graduate school trying to use the printed version of the Science Citation Index, I would have killed for a product like this. But it is just the first step. As stated above, the linkage to full text is a must. Even organizations that are both primary and secondary publishers (Elsevier, are you listening?) that

can link their full text to their own database should jump at the chance to link to everyone else's database. All that can happen is that more users will find their articles and want copies or access. And all you secondary publishers without full-text, what are you doing to provide links to full text? How many intermediaries do we need, especially when more and more bibliographic and abstracts databases are available for free?

There is a growing trend to link bibliographic databases and full-text [see sidebar]. The easy way is to simply allow the user to search through citations and abstracts and then link to the full-text. Some publishers take this a step further, by linking the references at the end of the article back to their database, and then to available full text. And there are a few examples of full text to full-text linkages, without moving through a bibliographic database in between. For now, the "forward in time" approach that Web of Science allows is rare. The Institute of Physics has done this with their own journals and the INSPEC database, but no individual publisher has the depth of coverage that ISI does. The biggest problem with linking cited references is the tremendous variation in the format of a cited reference -- hence the look-up feature for cited reference searching. It isn't just ISI that comes up with all those variations -- it's mainly the primary publishers. Automatically identifying a given reference is extremely difficult with such variation in formats. At the moment, ISI has the best solution to the problem. And it makes sense for publishers to use this expertise rather than trying to re-invent the wheel.

After full text, the next step involves going beyond subscriptions and journal-only information. I don't think ISI has seriously begun drinking this far into the future yet. Create databases that cover the entire spectrum of scientific communication. Include references and links to Web-based data, discussion lists, Web sites, and more. Like EBSCO Publishing's Collectanea and the Northern Light search service that search both proprietary collections and Web sites, IST should look at expanding database coverage to all kinds of scientific information. When it does, the product will truly deserve the name Web of Science.

Comparing the Online SciSearch Database

There are a few things you can do on the traditional online services that you can't do through the Web interface. The Web of Science has no equivalent to the Expand feature on Dialog. You cannot directly access the inverted index. Thus, you must account for spelling variations and cited reference variations on your own. However, the built-in lookup feature for cited references helps to get around this omission. Also, you can't Rank articles by most cited references in common. There are no sets, not that you will miss them, and you can edit a previous search and re-enter the revised search strategy. You can use up to 50 Boolean operators in one search statement, and separate entry boxes for different fields makes it easy to combine them in one search query. Though you cannot "de-dupe"

records at any point in the search, as you might with Dialog, all the records marked merge into a single "de-duped" set throughout the search session.

On the other hand, some features of the Web of Science you can't find anywhere else. First, the data is ideal for browsing -- something scientists love to do, but for which traditional databases don't provide. The interface and database are structured the way scientists think. They may do a general subject search, but usually prefer to find a key article or author and track that. They can look at the cited references, related records, or who cited that article, all with just a few clicks. That, combined with the incredible breadth of the database, makes it very suitable for serendipitous browsing or searching.

One example: I searched for the phrase "electronic publishing" through all three citation indexes, expecting most of the relevant articles to come from the information management type journals. Not so! The first article in my list of results came from the Canadian Journal of Plant Science and talked about authors submitting papers to electronic journals. Wow! A chance to hear from someone besides publishers and librarians (not that their views aren't valid), and hardly a journal I would have thought to check. Following the cited references gave me several more articles from scientific journals that provided an author's perspective to the future of scientific publishing. Too bad I'm still waiting for the full-text of the articles.

Lateral Searching and Citation Matching

Traditionally, users have searched databases using their own keywords to retrieve records. Browsing consisted mainly of reading through pages of citations, sorted in reverse chronological order. Today's technologies enable the user to move through the literature in completely new ways. ISI's citation indexes were always unique in allowing a user to follow the "trail of science" by finding papers that cite other papers, but their Web of Science implementation finally makes this easy. And, they have added another level of movement with the Related Records feature. This allows the user to find other articles that share at least one cited reference with any particular article, with just one mouse click.
Similarly, PubMed has built into the MEDLINE database a way to find related articles by using keywords. They created an algorithm that compares all of the words two documents have in common, with special weighting for title words, keywords, and MeSH headings, to find related articles. Again, all with one simple click for the user. And, of course, the MEDLINE records link to the full text from associated publishers using a unique PubMed number.

Other publishers have begun matching citations and full-text articles, too. The Institute of Physics uses the INSPEC database and Hypercite technology to create the linkages from one article's references to the full text, and even the "forward in time" notion of

who has cited a particular article of interest. BioMedNet has Citation Mapping between the full-text articles from their associated publishers and the MEDLINE database and also from the cited references to the database and back.

Now if only all these independent schemes interacted with one another, users could enjoy uninhibited browsing throughout the scientific literature, no matter where or how it is published.

The Linking Phenomenon

A Sampling of Linking Databases to Full Text

ChemPort http://www.chemport.org

This new service from the American Chemical Society only worked in demo mode at presstime, but expected to go live by the end of 1997. Users search the CA Plus database through STN or STN Easy (for which they must have an account). They pay search and display charges for records retrieved. Database records link to the full text on each publisher's site, and the user must have a subscription for each journal to gain access. Initially, Chemport will link to the 26 ACS journals, as well as selected journals from Academic Press, Springer-Verlag, Royal Society of Chemistry, and AAAS. ACS hopes to add links to additional publishers.

Elsevier's ScienceDirect http://www.sciencedirect.com

Currently Elsevier has the full text of about 90 life sciences journals available, and they hope to have all of their more than 1,000 journals available by mid- 1998. The full text links to the Embase database, so you can search either the database record or the full text of the journal or both together. Some of the journals only exist in an unsearchable PDF format, but Elsevier intends to put all the material up in HTML format.

Institute of Physics http://www.iop.org

IOP has put up the full text of all their journals and linked the articles and cited references to records from the INSPEC database. Their Hypercite technology enables you to follow links to references cited in an article (from 1996 forward), or papers from institute of Physics Publishing's archive which cite the current article. This is achieved through links to INSPEC's database dating back to 1969, the Los Alamos preprint server, and Institute of Physics Publishing's own abstract and full text archive (1996-1997). You must have a subscription to each journal for which you want access. IOP hopes to develop additional cross-publisher linking with other physics publishers and pre-print services.

OCLC Electronic Collections Online

http://www.oclc.org/oclc/menu/eco.htm

At the moment, ECO operates separately from OCLC's patron-oriented FirstSearch service, but OCLC hopes to integrate the two in 1998. At first, users will be able to search and browse ECO from within Firstsearch, retrieve abstracts and articles for the journals their libraries carry, and purchase individual articles for journals to which their library does not subscribe. Eventually, OCLC will link FirstSearch databases that index articles in Electronic Collections Online to the full-text articles.

In addition, OCLC has contracted archival rights to the journals in Electronic Collections Online. OCLC maintains a subscription profile for each subscriber's journals. If the subscriber maintains an active Electronic Collections Online Access Account, OCLC provides ongoing access to the data published while each journal subscription remained active. A library's rights to the archive remain active regardless of current journal subscription status. OCLC's FirstSearch and EPIC online services can indicate library holdings.

PubMed http://www.ncbi.nlm.nih.gov/PubMed

Produced by a separate unit of NLM from the Medlars group, this site provides free access to the entire MEDLINE database and links to the full text of several associated publishers, currently about 95 journals. The links take you to the publisher's site, where you usually need to have a subscription to access the full text. PubMed also provides links to related articles, as determined through an analysis of keywords and MESH terms.

BioMedNet http://www.biomednet.com

Access to the MEDLINE database is provided through free registration with BioMedNet. The database records link to the full text of associated publishers. You must set up an account with BioMedNet to access the full text.

When the corresponding full-text paper for a MEDLINE record exists on BioMedNet, you can link directly in both directions -- from the article to the MEDLINE record and vice versa. But with Citationweb you can also follow links from a MEDLINE record to all the articles on BioMedNet citing the item. They hope to extend this system soon to journals outside BioMedNet.

Medscape http://www.medscape.com

Similar to Biomednet, except this service only puts up the full text for selected articles from associated publishers. The full-text articles contain links to other related Medscape articles, other related Web sites, and an annotated MEDLINE search.

Ovid http://www.ovid.com

Ovid currently offers the full text of over 80 biomedical, mental health, and nursing journals, soon to be over 350 journals (from more than 40 publishers) as part of six separate collections with references linked to the most pertinent bibliographic databases, like CINAHL, MEDLINE, and PsycInfo.

On a pay-as you-go basis, users can access the full-text databases directly, or follow links from MEDLDM or one of the other pertinent databases and access and pay for individual records from the full-text collection. Of course subscription-based pricing is also available. The journals store in SGML format, which allows for fielded searching within full-text articles. Ovid can also indicate library holdings on records.

Blackwells' Electronic Journal Navigator
http://navigator.blackwell.com

This source links to several publishers' sites to provide libraries with a single point of access to their electronic journal subscriptions. The service does not currently link to any database, but since the Web site also provides links to database services from Cambridge Scientific Abstracts, H.W. Wilson, UMI, ISI, and others, it appears that they hope to link access in the future.

Science Magazine Online http://www.sciencemag.org

Science not only links their cited references to the MEDLINE database through PubMed, but also to the full text of articles in other electronic journals put online through High Wire Press. in many cases you do not need a subscription to see the linked full-text article, but if you try to explore other articles in the linked journal, the system wants your password.

EBSCO Publishing's Collectanea http://www.epnet.com

Although primarily business oriented, this site provides access to full-text articles from over 1,600 publications and links to over 15,000 Web sites for a flat fee. Full-text articles contain hyperlinks to company profile pages with contact information links to company descriptions and other related topics.

SearcherFeb, 1998

Webmaster liability: look before you link, and other admonitions for today's Webmaster.(includes related articles on tips and resources Webmasters could use to protect themselves against liability lawsuits)(Sidebar)

Author/s: Carol Ebbinghouse

So, you have a Web page. Congratulations! People now have yet another way to reach you, to learn about you, your family and/or your company. They can do this at times when you remain unavailable by telephone, fax, beeper, or cell phone -- and they don't even have to wait for you to respond to your e-mail! Your company can put out all the information describing its products, services, professional qualifications, and any other information (ordering, delivery, terms and conditions, etc.) that will help, inform, entertain, and even extract money from the reader.

On your personal or family page, you probably have photographs of you or the kids, a cartoon or two, maybe even a cute little Java application or applet. Your Web page may provide links to other informative, educational, humorous, or related Web sites. You may allow guests to post messages to you or to the Web page itself. It's fun, as long as you are not the Webmaster with responsibilities for creating, keying, and updating all the information, locating and maintaining the accuracy of links, and preventing the any number of things that can go wrong.

Oh, sorry, you ARE the Webmaster? The one responsible for constantly upgrading the site to keep it interesting? The one who keeps adding and checking and revising the link addresses? The one constantly on the lookout for funny items, newer pictures, more and better organized links? The one who reads all the postings from readers? Gee, that's too bad. Why? Because it is a lot of work and a thankless job? No, because on top of all the other chores belonging to your status as a Webmaster, now you can also get sued -- maybe even arrested!

Lawsuit? Liability? Arrested? Because of a Web page?

You betcha. Civil lawsuits against Webmasters can be brought for a number of activities, and criminal prosecutions are happening as well. It may not even matter that your Web site is legal where you live. Lawsuits and criminal prosecutions are being pursued in some state and federal courts for activities in other states.

The Federal Government

Big Brother is watching the Web in an effort to protect consumers.

Ads, claims, and services sold on Internet Web sites can raise legal issues in a number of ways. For one thing, the Federal Trade Commission (FTC) conducts "surf days" to discover Web sites that make claims and contain promotions for products and services that mislead consumers The FTC's Internet Coordinating Committee regularly conducts "surf days" for scams in the work-at-home market, fix-your-credit, college scholarship scams, and travel fraud, in addition to responding to complaints about fraudulent schemes coming from consumer groups. Late in 1997, they focused on health claims relating to treatments and "cures" for AIDS, arthritis, cancer, diabetes, heart disease, and multiple sclerosis. According to the FTC, roughly 25 percent of the sites found on "surf days" that receive FTC warnings just shut their sites down before formal investigations even begin.
The FTC also has set up a variety of Web pages that impersonate con artist promotions on the Internet. Deliberately luring unwitting consumers into home business and marketing schemes with juicy come-ons and testimonials, they ultimately hit visitors with a "YOU COULD GET SCAMMED" warning about fraud on the Internet.[1] (Visit the Federal Trade Commission at http://www.ftc.gov.)

Recently one Webmaster started using the Net to canvas for investors for his revolutionary new toothbrush enterprise. Self-posting an initial public offering runs afoul of both state and federal requirements regarding registration of securities. This individual now faces state and federal investigations with civil and criminal consequences. In addition, AOL (America Online) has reportedly cut his Internet access off in light of the investigations.

The Securities and Exchange Commission (SEC) has a whole division investigating investment fraud. (See http://www.sec.gov/enforce/litigrel/lr15237.txt and/or http://www.sec.gov/enforce/litigrel/lrl5490.txt.) So do other federal regulatory bodies, e.g., the Commodity Futures Trading Commission (http://www.cftc.gov/). When someone runs afoul of federal securities laws, state enforcers follow close behind in investigating and imposing sanctions, which can include disgorging all profits, civil and criminal fines, and even imprisonment. And state enforcement actions can come from every state where the securities are offered!

Of course, one ought not do on the Internet that which is illegal in print, either. An Internet newsletter writer, Theodore Melcher Jr., pleaded guilty to stock fraud and violations of federal securities and tax laws. In September 1997, a federal judge sentenced him to a year in prison and fined him $20,000. Mr. Melcher admitted to using his newsletter, the SGA Goldstar Whisper Stocks, to issue false favorable press releases about Systems of Excellence to drive up the value of his shares to an inflated market price.

Finally, as Virgin Atlantic Airways learned in February 1996, the federal government cares about whether Web pages in highly regulated industries stay up-to-date. Based on two misleading statements in the fare list on its Web page, Virgin Atlantic was fined $14,000 by the Department of Transportation for false advertising. While valid when posted, the information on the fares had changed by the time potential customers called to book flights. The federal agencies obviously consider that truth in advertising requirements apply to the Web, just as they apply to print, radio, and television.

The States

Web owners have more than the federal government to worry about. The states of Wisconsin, Minnesota, and Missouri are suing Idaho's Coeur d'Alene Indian Tribe for their "U.S. Lottery" game (http://www.uslottery. com) on the Internet because gambling is illegal in those states. They are bringing these actions even though gambling is legal in Idaho and being conducted on a U.S. Indian reservation. Internet gamblers set up accounts with credit cards to pay gambling debts and receive winnings. The virtual gambling sites do not operate under the same strict controls relating to ownership, financial reserves, or fairness as regulated casinos. The same states have also begun lobbying for a proposed federal Internet Gambling Prohibition Act, which would provide for fines and jail terms for even casual gamblers. A variety of bills relating to Internet gambling, including provisions allowing state attorneys general to file federal court actions, are currently under consideration in Congress.
States can have long arms. In another case, a bulletin board system (BBS) sysop[2] and his wife in Milpitas, California, were dragged to the western district of Tennessee to defend themselves against charges of interstate transportation of obscene materials. Their BBS, called "the nastiest place on earth," provided Graphic Interchange Format (GIF) files of photos and films of nudity, bestiality, and sundry other specific types of nastiness.[3] Tennessee tried them under the local community standards of western Tennessee (not California where they lived and housed their computer and files). And they were convicted.

What's in a Name?

Getting a domain name does not create a trade name; however, using someone's registered trademark in your domain name will offend the trademark owner. Hence, before you dream up your

domain name, check trademark databases or hire someone to make a trademark search to verify that your domain name will not confuse anyone, or get you in trouble with the trademark holder. Then get a trademark of your own. For example, an entertainment spot in Missouri called the Blue Note set up a Web page which a New York City-based Blue Note sued in federal court in New York. The court threw the case out because of insufficient contact to give a New York-based federal court jurisdiction. The Missouri Blue Note Web page had no appeal to New York consumers, was directed to a local audience, received no money from New York sales, and had no physical presence in New York. However, the court did not seem to rely upon the disclaimer on the Missouri Blue Note's Web page that readers should not confuse it with any other club operating under the Blue Note name. Undaunted, the owners of the New York Blue Note indicated that they might go to federal court in Missouri, which would have jurisdiction over the Missouri Blue Note, to protect its name. (Bensusan Restaurant corp. v. King; see http://www.law.vill.edu/ fedct/fedcourt.html). Pity the Web owners who must pay to defend actions in every court in which they are prosecuted.

Also, do not try to put one over on the government. An allegedly pornographic board that had secured the "nasa.com" and other "nasa" domain names was forced to change its name after people trying to locate the Mars Pathfinder photos (http://www.nasa. gov) got something else entirely!

Potentially even more formidable than the U.S. government is Microsoft Corporation. Ask Daniel Khoshnood, a Cal-State Northridge student, who called his Web design company's Web page "microsoftnetwork.com" and now has the real Microsoft Corporation coming down on him for copyright infringement, trademark infringement, and unfair trade practices. Apparently Microsoft called Mr. Khoshood's Internet service provider (ISP), who shut down Mr. Khoshnood's access, believing that his business wasn't worth the grief Microsoft could cause.

The Federal Trademark Dilution Act, passed in 1995 (15 U.S.C. [sections] 1127), has appeared in several cases in its short life to protect famous trademarks. Hasbro Inc.[4] successfully protected its Candy Land children's game when it obtained an injunction against an "adult entertainment" Web site that had obtained the domain name "candyland.com." In another adult entertainment case, the "Adults R Us" domain name was enjoined after Toys "R" Us, Inc.[5] sued. In a more political case, an avid proponent of the anti-abortion movement acquired the domain name "plannedparenthood.com" to proclaim a pro-life message. Planned Parenthood Federation of America, Inc. sued[6] and won.

In a series of cases, several corporations have sued an enterprising alleged "cyber squatter" named Dennis Toeppen. Toeppen has apparently acquired over 200 domain names using trademarks (including "eddiebauer. com," "nieman-marcus.com," "anaheim

stadium.com," "frenchopen.com," and "aircanada.com," among others). In one case, Mr. Toeppen allegedly offered Panavision the domain name panavision com for $13,000.[7] In another case, Intermatic Inc.[8] sued Mr. Toeppen to enjoin his use of "intermatic.com" as a domain name.

Even if your domain name does not infringe or dilute a trademark or service mark, you can still run afoul of trademark owners if you use "meta-tags" data to lure unsuspecting consumers to your Web page.[9] In one case, a pornographic Web page did two things to offend Playboy Enterprises, Inc. First it used the trademarked names "playboy" and "playmate" in their domain names ("playboyxxx.com" and "playmatelive. com"). Second, the "meta tags" for those sites used the trademark terms so many times that a search of "playboy" on an Internet search engine ranked the porn site as number three for relevance.

If You Can't Say Something Nice...

Disparagement of a trade name can also be a no-no. In one incident, Presstek Inc. sued three people for allegedly defamatory and inaccurate statements made about the company in online chat groups. The suit claimed that three short-sellers (who would profit from a decline in the price of a stock) made false statements about Presstek, alluded to a possible grand jury probe, predicted a loss for a quarter that turned out profitable, and reported that the auditor was about to resign over revenue-recognition issues. This is but one of a number of risks associated with chat rooms and other content posted by users or visitors to a Web site. Businesses can and will sue Webmasters to protect a good name.

Finally, as Matt Drudge of the Drudge Report on America Online and the World Wide Web can attest, you should never say anything nasty about someone unless you are absolutely convinced of the truth and newsworthiness of the statement. Mr. Drudge now faces a $30 million libel lawsuit brought by Sidney Blumenthal (an aide to President Clinton) because of a false report that Mr. Blumenthal had a "spousal abuse past that had been effectively covered up." The quick retraction of the false statement only mitigates damages, it wouldn't eliminate them. This case will likely be in litigation for years, because it raises a number of issues in relation to the Internet and established print-based libel laws.

Intellectual Property

Copyright issues always arise when you load text, photos, cartoons, sound or film clips, and other creative pieces. The simple rule of thumb is to take your own digital photos or scan in pictures that you took yourself and have a signed model release on file from anyone featured in the photo. Write your own text, and don't post anything like comics, photos, or documents without a release from the creator. In my opinion, this constitutes the most fertile area for

Webmasters to be held personally liable -- with civil and criminal consequences. The Webmaster should neither post copyrighted works, nor permit reposting by visitors. Public domain material can be copied not because it is public (e.g., on the Web), but because it was created so long ago that it is no longer protected by copyright. Project Gutenberg has built a wealth of materials in the public domain, and placed it all on the Web.[10] For safety's sake, assume that anything posted on the Web (public or not) is still covered by copyright, and that authors, publishers, or other owners of the copyright will enforce their rights through civil and criminal actions. A group called Webposse, composed of hundreds of photographers, publishers, and artists, fights illegal use and exploitation of members' work. They allege that free copying of their work destroys their ability to earn a living from the fruits of their creative labor (http://www.Webposse.com/index/htm). The group has two main objectives: to educate the public on intellectual property issues and to protect their creative works from copyright violators. The group is also investigating an electronic "tag" that will enable them to quickly search for copies of their work. Webposse is not the only group performing vigilantly for their constituents. The Software Publishers Association has a massive program in place to prevent and punish freeloaders. The Writers Guild defends their members' interests, too.

In addition, a new industry has begun to sleuth out pirated software throughout the Internet on behalf of the software developers, producers, and publishers. A new breed of Internet detectives have formed companies now being hired by Hollywood entertainment companies and the music industry, in addition to hundreds of companies seeking to protect their trademarks. With names like Infringatek, Inc., Markwatch, and Datalytics, Inc., these companies reportedly do a booming business in a growing industry.

Viacom, Inc., owner of the Star Trek brand, has stepped up Internet enforcement of its rights and sent "cease and desist" letters to Webmasters of Star Trek fan sites that post copyrighted film clips, sounds, and emblematics. Viacom has posted an official statement outlining its Internet enforcement policy (http://www. paramount.com/openletter/) and encourages people to report unauthorized postings of their intellectual property on the Internet (much like a neighborhood watch).

Webmasters; should never post anything for copying unless it is their own work, definitely in the public domain, or they have clear authorization to post it. Furthermore, they should vigilantly insure that no one else posts copyrighted music, software, documents, photos, cartoons, or other intellectual property on their Web site without permission from the author or owner. New federal legislation is being proposed in the copyright area, relating to computers, fair use, and other issues.

Terms and Conditions

Finally, Webmasters, protect your own work--especially, on your own Web site! Use the copyright symbol on your work and if you don't want it copied, say so.

If you allow others to post to your area, then post a "terms and conditions" agreement insisting that those who would post to your site will follow copyright and other intellectual property law (only post their own original work, etc.). If you want to further limit your liability, then consider posting a disclaimer for what others upload or say in chat and other areas. If you want to post disclaimers, then do so, with a contract that both limits your liability and assures that the individual posting illegal copies or offending material will indemnify you for any damages you may suffer because of their posting. But beware that "click here to accept the terms" contracts are not valid in all states.
Terms of use (also known as "terms and conditions") on some sites can make people posting material to your Web page liable, not just to those defamed or injured, but also to you, the Web site publisher, under indemnification clauses. Terms and conditions specify user rights and responsibilities and may or may not require a "click to accept" the terms. Not all jurisdictions will enforce "click to accept" contracts. The harder to find and more buried the contract is, the less likely it is to be enforced. If you don't require a "click," then the chances of enforcement become even less. Some users look on "click to accept" contracts as merely a form of intimidation, but few will want to spend their life savings to litigate the issue. (See Warner Brothers Online "terms of use" at http://www.warner bros.com/terms.html and sites listed in Ebbinghouse and Arditto, "All Rights Reserved. Well, Maybe Not: Copyright in a Web World," Searcher, vol. 5, May 1997, pp. 24+.).[11]

If your Web does not allow postings, then this should not affect you. However, if you do allow postings or chats, a disclaimer and/or term of use might help to dissuade or deter those who might post defamatory or copyrighted or trademark protected pieces. Should you decide to use an agreement, especially an indemnification clause, then post terms and conditions up front on the Web home page. Make sure every new user, and most importantly anyone trying to post material to your site, must click an assent to the agreement in order to get past it and into your site.

Kraft, a subsidiary of Philip Morris companies in New York, has "a message from our lawyers" at the bottom of its home page. These terms and conditions warn visitors against posting copyrighted recipes, notes limited liability for any damages resulting from the use of the Web site or the recipes that it contains, etc. (http://www.kraftfoods.com). This example shows a clearly written terms and conditions page easily apparent from the home page, but this does not make reading the terms and conditions page a pre-requisite to exploring the site itself.

Linking and Blocking Links

Linking from one site to any other constitutes one of the best features, indeed the backbone, of the World Wide Web. Finding a site with great links and bookmarking it lets you come back time and again for the best and latest links on a certain subject, without clogging your "favorites" or "bookmark" file with more than it can handle. A Web site with good links creates value and good will on the part of users, who can become loyal supporters and users of your Web page. If you have advertisers on your page, they can profit from the popularity of your many and well-organized links that draw new users to your Web page all the time.

Some Web pages require licenses for other Web pages to link to their page. Other sites only want you to ask for their permission (to make sure that you are not a competitor, doing anything illegal, etc.). Check the terms and conditions on any site to which you would like to create a link. If still unsure, check the site for the e-mail address of the Webmaster and write a letter.
TotalNews.com had been creating links to other sites, including commercial sites, while maintaining a frame with its own advertising superimposed around the frame with the linked information screens. Dow Jones (publisher of The Wall Street Journal), Times Mirror, Washington Post, Reuters News Media Inc., and others sued TotalNews based on copyright and trademark issues. Basically, the suit alleged that TotalNews had made the plaintiff's information look like its own (surrounded by TotalNews' frame and advertising) and deprived plaintiffs and their advertisers of credit, visibility, and revenue. The lawsuit resulted in TotalNews settling and agreeing to stop framing the plaintiffs' sites.[12] A similar case arose in Scotland between two local newspapers. The Shetland Times Ltd. sued the Shetland News alleging that the links between the News' site and The Times main headline page made it appear that the News created the articles. According to ComputerWorld[13] a judge granted an "interim interdict" blocking the links pending a full hearing on the case.

Microsoft began linking from its Seattle Sidewalk local community page to Ticketmaster's ticket order screen for users to order tickets to local events. While Ticketmaster and Microsoft had negotiated over revenues stemming from such a link, Microsoft just began to go ahead and link to the Ticketmaster order screen, bypassing the Ticketmaster home page and its advertisers. Ticketmaster blocked links to the order screen from the Sidewalk, referring users instead to their home page, and sued Microsoft. Since then, Ticketmaster has allowed direct links to its site from CitySearch and has announced a deal with Excite, so users of that service can link from discussions of an event to Ticketmaster and order tickets to the event.

Soon some sites may charge for a license to link. To determine what Netiquette rules apply in establishing a link to a site, "click" on terms

and conditions. In the case of "a word from our lawyers" on the Kraft Foods unit of Philip Morris Companies' Web, the site asks visitors not to create any kind of hyperlink from another site without "written consent -- first."

You could get into trouble even if you have permission to link to a home page. If your Web links bypass registration or payment screens on the home pages of sites to which you link, you may create trouble for yourself and roadblocks for your users. Many sites request registration or marketing information from visitors; others have advertisers or charge for access. You will anger the Webmasters of such sites if your links take users past these screens and paid advertising directly to the good stuff. This is what got Microsoft in such trouble with Ticketmaster, and "framing out" the advertisers of linked sites was the root of TotalNews' problems.

Finally, protect yourself: Make sure that the sites to which you link are legal in your own state and in others. Be wary of linking to gambling sites, lottery sites, pornography, and any other sites that could get you in trouble -- even if your state deems the activity legal. Think before you link.

Foreign Trouble

Can you run afoul of laws in other countries? Only a few incidents have come to my attention, but they indicate that Webmasters might have problems in this arena as well. Germany brought a CompuServe executive up on charges of not doing enough to block pornography and Nazi-oriented materials originating in the United States from coming into Germany. Germany has some of the stricter laws on Internet content. A student named Angela Marquardt was prosecuted for a link from her home page to Radikal magazine, which provides articles on terrorism, including an article on how to derail trains. The magazine resided on the Internet service provider xs4all in the Netherlands. Charges were later dismissed because her link was created before the article on derailing trains was actually posted. Prosecutors were considering an appeal.

In Quebec, Canada, a Webmaster was threatened on May 29 with a fine of $1,000 for having an English-only Web page in a French-speaking province. The Web site from Micro-Bytes Logiciels had a little over a month to provide a French version of the Web site before the court would impose the fine (http:// www.microbytes.com/main.html).

In a similar prosecution, this one by the French Ministry of Culture, Georgia Tech Lorraine (the European campus of the Georgia Institute of Technology based in Metz, France) was sued for having an English-only Web site. The site has instituted French and German versions of its Web site [http://www. georgiatech-metz.fr/], though the court had dismissed the charges on a technicality that plaintiffs failed to file a police complaint before pursuing civil suits. Similar

charges were filed against The Body Shop last year and it lost.

Conclusion

While most Webmasters sail through life without incident, we must keep in mind the potential for liability. Many Webmasters load disclaimers[14] of liability -- especially for postings by others to their site. Other Webmasters add terms and conditions, which users may or may not have to "click to accept" in order to gain access to the Web pages.

In light of the potential for liability, Webmasters should carefully and regularly reconsider their links, review their use of trade names, verify the copyright ownership of text, graphic, and photographic material on their site, and check the tenor of postings that may disparage individuals or trade names or interfere with business relationships.

Perhaps before you start a Web page of your own, you might want to look into a personal liability rider on your homeowner's insurance policy, or your business might look into some of the new insurance plans offered for Internet activities.[15] If you get a "cease and desist" letter from a law office, or are served with a complaint in a lawsuit, or face criminal charges, you should see an attorney that specializes in the field right away. Just remember that even though you might prevail in a civil or criminal action, but the process itself could exhaust your personal resources.

[1] See Kathy M Kristof, "Personal Finance: Surfing for Scams" Los Angeles Times, Sunday, November 30, 1997, p. D2. Also see "FTC Warns a Slew of Sites About False Health Claims," Wall Street Journal Interactive Edition, November 5, 1997, and http://www.ftc.gov.

[2] I don't know how many of you remember when you had to dial into someone's home computer to chat and gather information, but BBS systems still exist, and I believe BBS case decisions could easily extend to Webmasters.

[3] See United States v. Thomas, 74 F. 3d 701 (6th Cir. 1996).

[4] See 40 US.P.Q.2d 1479 (W.D. Wash. 1996).

[5] Toys "R" Us, Inc. v. Akkaoui, 40 U.S.P.Q.2d 1836 (N.D. Cal. 1996).

[6] See Planned Parenthood Federation of America, Inc. v. Bucci, 97 Civ. 0629 (KMW), 1997 WL 133313 (S.D. N.Y. March 24, 1997).

[7] Panavision International, L.P. v. Toeppen, 945 F.Supp. 1296 (C.D. Cal. 1996).

[8] Intermatic Inc. v. Toeppen, 947 R Supp. 1227, 1996 WL 716892 (N.D.Ill. 1996).

[9] See Carl S. Kaplan, "CyberLaw Journal: Legal Roadblocks Starting to Deter MetaTag Hijackings," "CyberTimes The New York Times on the Web" October 16, 1997. Also look at Playboy Enterprises, Inc. v. Calvin Designer Label the preliminary injunction is at http://www.jmls.edu/cyber/ cases/calvinl.html, and the amended complaint is at http://www.patents.com/ac/ playcpt.sht. Similar issues were involved in litigation involving Oppedahl & Larson against Advanced Concepts in Colorado; see http://www.patents.com/ac. Another case resulted in a consent order against a firm using plaintiff's trademarks in meta tags; see http://www.cll.com/wnewfr.htm. The Excite search engine ignores meta tags in its indexing and now filters to cut down on the abuse of hidden keywords and phrases in Web sites. See http://www.excite.com/Info/listing. html#anchor4877066.

[10] See Project Gutenberg at http://www-promo.net/pg/.

[11] Other examples of terms and conditions on the Web include: AllPolitics Service Agreement at http://allpoliticscom/1997/ utilities/info/terms.htmi, or the New York Times on the Web subscriber Agreement at http://www.nytimes.com/subscribe/help/ agree.html and the Nando Times User Agreement at http://www.nando.net/nt/usra gr.html.

[12] See http://wwwtotalnews.com/.

[13] Mitch Wagner, "Internet Commerce: Web Firms Eye Suit on Link Policies," ComputerWorld, May 12,1997, pp. 61+. See also http://www.computerworld.com.

[14] For examples, see the sidebar "Resources for Webmasters."

[15] The Chubb Group of Insurance Companies made headlines in August 1997 by offering insurance to protect Internet-content companies and Internet-service providers that inadvertently misuse copyrighted material or allegedly defame a person or business. See http://www.chubb.com.

RELATED ARTICLE: Resources for Webmasters

The following sites and cites can help you stay on top of the changing legal environment for Webmasters. They can help you stay out of trouble, or know when you've landed in it.
Internet Legal Survival Guide

http://www.gtlaw.com.au

The law firm of Gilbert & Tobin has prepared publications on Internet

compliance and an Internet Legal Survival Guide. For a good article on Webmaster Liability, try http://www.gtlaw.com.au/gt/bin/frameup.cgi/gt/pubs/ survival guide.html. For the list of publications, try http://www.gtlaw.com. au/cgibin/wwwwais.bas?keywords=gtpublication&gttopics=internet

Intellectual Property Law Web Server

http://www.patents. com

The Intellectual Property Law Web Server, provided by the law firm of Oppedahl & Larson, has a library of Web publications as well.

QuickForms

http://www.quickforms.com

The QuickForm contracts Online site for their Online Service Provider Agreements, Internet Advertising Agreement, Web Site Development/Hosting Agreement, and Legal Survival Guide for Computer Industry and Internet Commerce Transactions.

Other Cyberspace Law courses

http://www.jmls.edulcyberlothers.html

Cyberspace Law Institute

http://www.cli.org

They also offer a sysop liability course at http://www.cli.org/sysopl/ sysopl.html.

Legal Webmasters List

Legal Webmasters have their own listserv! To subscribe, send an e-mail with no subject to listserv@listserv.law.cronell.edu with the message: subscribe LEGAL-WebMASTERS yourname. To send postings, use legal-Webmasters @listserv. law.comell.edu.

The LEGAL-WebMASTERS list also pen-nits people to post messages in daily digests if you send the following message: To listserv@listserv.law. cornell.edu, set LEGAL-WebMASTERS mail digest. Try this command for any listserv that swamps you with reams of daily messages! It works.

[Personal note: I believe that digesting listserv input/output has saved my e-mail box from overflowing and me from hours of time spent reading irrelevant messages and from the dangers of blindly deleting messages to fit everything into my electronic mailbox. I

recommend digesting for listserv subscriptions (especially unmoderated listservs).]

RELATED ARTICLE: Webmaster Do's and Don'ts

Protect your domain name. Do a Trademark search before selecting your domain name. Avoid using any trademarked name. Then, trademark your domain name if you have a business interest in creating it (see http://www. uspto.gov/Web/offices/tac/doc/basic). Don't create metatags with trade names unless they accurately indicate the purpose or mission of your Web page. In short, do not "pack" trade names to lure people looking for the real trademark owner's site. Vigilantly search out others who begin abusing your trade/domain name and call them on it. If your trademark/domain name affects vital issues, like your income, you might consider using MarkWatch or some similar watchdog service mentioned in this article and in the "Contacts" (pages 69-72) section.

Protect your own work first. Put copyright symbols on each page of your Web site. You have created something, and it is worth protecting. You might want to consider registering with the copyright office (http://lcWeb.loc. gov/copyright/), since registration provides enhanced protection in infringement suits (such as statutory damages and attorney fees). Don't hesitate to use technology to protect your site: firewalls, encryption envelopes, metering, access codes, digital certificates, dynamic pages, and other tools discussed in the literature can all help.

Protect the work of others. Do not post the work of others without their permission. That goes for copyrighted material, trade names, trade secrets, software, text, photos, cartoons, and other illustrations. When you have permission to copy the work of others, note that fact on the screen, and include the copyright sign and the creator's name. Make it obvious that you take intellectual property laws seriously. Post only photos that you take yourself and get permission from all people in any picture. If someone notifies you that you have posted infringing, defamatory, or libelous material on your site, remove it immediately if the claim is valid.

Look before you link. Do you need to ask for permission? Courtesy might dictate that you approach private or personal Web sites for permission to link. Also contact those that require membership or registration, carry advertisers, or otherwise indicate an interest in who visits their site. Link to the initial screen -- not a buried location -- or wherever the permission would dictate. State and federal government information sites as well as academic and library sites would probably rather have visitors than limit the links, but even with these sites it would be Netiquette courtesy to ask.

Watch who links to you. Keep track of sites that link to your pages. Scripts can be written that notify you when someone links to your site. Unsuitable links may tarnish your image. Be prepared to write a letter to any offending Webmaster in an effort to resolve the

problem without the need for an attorney. Some pages have a notice to anyone wishing to link to the site, with instructions on how and where to link and/or admonitions against certain kinds of links.

Proclaim the disclaimer (and don't forget the indemnification clause). Alert users to your copyright and trademark interests, link policy, and limitations on unauthorized use and posting. Don't forget to include an indemnification clause in the event you are sued for the behavior of visitors that post to your site. Carefully consider the potential liability for defamation and other actions should you allow posting by visitors to bulletin boards, chat area, guest book, or comment sections. Guests can potentially load defamatory or pornographic, misleading or deceptive material, descriptions of the means to commit violent crimes, etc. Recent headlines all urge caution.

RELATED ARTICLE: How to Find a High-Tech Attorney

So you didn't get this article in time and you've gotten into trouble. Or maybe you walked into trouble with your eyes open. Maybe you have chosen to fight. Either way, it's time to get a lawyer -- and the right lawyer.

Assuming you have access to online resources, check lawyer directories for attorneys who limit their practice to high-tech, copyright law in particular, intellectual property in general, etc.

Check the largest directories of attorneys -- MartindaleHubbell, available on Lexis-Nexis (MARHUB) and on the Web (http://www.martindale.com/locator), and/or West's Legal Directory, available on Westlaw (WLD) and on the Web at http://www.wld.com. The West Legal Directory includes an article on how to find an attorney and an explanation of different fee structures, among other things.
You can link to some of the most helpful Internet directories through the Washburn University School of Law's WashLawWEB at http://lawlib.wuacc.edu/washlaw/direc try/directry.html and http://lawlib.wuacc.edu/washlaw/ directry/directry.html.

Consult other directories of attorneys at the local law library run by regional bar associations (state, county, and city) or university law school libraries. Local law libraries should have all kinds of regional and specialty law directories.

Call your local bar association and ask if it makes referrals by specialty, or whether they can recommend a referral service. Many city, county, and state bar associations have a Web presence. For a quick list of those with Web sites, visit http://lawlib.wuacc.edu/washlaw/bar.htmI and http://lawlib. wuacc.edu/washlaw/bar.html.

Go online and search high-tech business databases for news on legal disputes or legal databases for articles on high-tech issues. If the files have fields listing where authors work, you can easily find attorney-authors located near you. You may not only get the name of an attorney well-versed in the subject matter, but also pick up some information on the issues and legal terms for the matter you will discuss with the attorney you choose. Look up cases (on legal databases such as LEXIS-NEXIS or Westlaw, or on the Internet) on your high-tech issue to see who represented whom, and who won. If the attorney won't take your case, perhaps s/he will recommend someone equally or more qualified. The Internet is full of law firms with Web pages, and the more high-tech the firm's expertise, the more likely they are already on the Net.

The author believes that personal recommendations constitute the best way to find attorneys and other professionals. Feel free to call InterNic (http://www.internic.com) at Network Solutions Inc. if you have a Web page domain name issue. Sometimes they will share the names of attorneys they know. Other organizations interested in "cyberlaw," e.g., the Electronic Freedom Foundation, might also have some names. Law school faculty who teach in the area of copyright, intellectual property, high tech (or cyberlaw), etc., often know key members of the local bar in their fields.

Do not forget Webmaster and computer user groups. You are probably not the first person in these circles to have your particular legal problem, and someone else may have already "won" their case and can recommend an attorney. (If they lost, maybe they remember the name of opposing counsel.)

You can visit personal and/or corporate Internet sites that list "computer law" attorneys, e.g., http://www.notice. com/lawyerfind.html; http://www.notice.com/lawyerfind. html; http://www.haas.berkeley.edu/~wehrli/; http://www. kuesterlaw.com; or http://www.eaglelink.com/power/ list.html. These lists may or may not reveal the criteria for inclusion, i.e., whether their status as highly respected individual attorneys and firms or fees paid lay behind the entry. Washburn University and other law school Web sites will steer you to bar-certified attorney referral services and other reputable sources with clear reasons for why and how they include attorneys and the basis of their "specialization."

Whatever you do, do not ask your family attorney for high-tech legal advice. Ask for a referral maybe, but not for advice. This is a specialized and fast-moving field with new laws, new regulations, and new court cases coming up all the time. General practitioners may be interested in this field, but you need a specialist.

Once you have decided on an attorney or two to check out, see whether your state bar association has a Web page with information on specific attorney members. In California, the State Bar Association [http://www. calbar.org] has a directory of members,

gives bar numbers, and tells searchers whether an attorney is an active member in good standing or not

Good luck!

SearcherFeb, 1998

Wins and losses: the latest developments in online competitor research.

Author/s: Helene Kassler

Marriages, wars, births, deaths, mergers, acquisitions -- a Shakespearean tragedy? A Wagnerian opera? No, just 1997 and another year of tsunami-scale changes in the online world with impacts touching many areas of competitor intelligence (CI) research.

Last year's changes clearly spotlight a transformation in the online industry. I -- and everyone else -- have seen the future and it is the Internet. At Fuld & Co., a consulting firm specializing in competitor intelligence, we find ourselves increasingly relying on information gleaned from both free and fee-based Internet services. One example of a free resource: A client retained us to determine the technological base for a competitor's electronic commerce Web site. My search on Monsterboard, the commercial job Web site (free to searchers), displayed more than 100 job positions open at that firm -- revealing vast details about the competitor's hardware and software use.

At my company, we even pay for Internet-based information services, when they offer value. Several project managers at Fuld subscribe to an affordable Internet-based, direct-to-desktop alerting service. Some profiles retrieve industry headlines; others specify company or product names. All the service's users find value in news, filtered to their needs, delivered to their desktop in a timely fashion.

It seems clear that the most exciting and useful services and changes are now developing on the Internet. While the older, mature companies port their massive databases over to the Internet with few Internet-enhanced innovations, in contrast, companies building new services from the Internet ground up (usually newer, smaller firms) often offer greater functionality by taking full advantage of the Internet's innate abilities such as hyperlinking,

graphics, and interactivity. Even within the Internet world, the rate of development and innovation is staggering. The game of competitive leapfrog takes place in a matter of months. One Internet moment finds a particular service the darling; three months later, a brand new service emerges with advanced features, greater content, a simpler interface, or a lower cost.

While thrilling, the Internet world can also be challenging and exasperating. Thrilling in its new services, capabilities, and usually low costs; challenging to keep up with new resources; and exasperating when one finds a favorite Web site redesigned with a favorite function buried five mouse clicks deep, or worse the site has vanished, or even -- heaven forbid -- now costs real money!

Fresh New Faces

In the commercial online world, innovations offering the greatest value to CI researchers have arisen from creative new systems or upgraded services from companies such as IAC, Manning & Napier, and Northern Light. Among the big three traditionals (Dialog, LEXIS-NEXIS, Dow Jones), only the new Dow Jones Interactive service utilizes the Internet's native talents at all well, and even there, restricts its "Netted" offerings to a special section featuring business Web sites. The Web's natural functions have not yet been completely integrated into the actual Dow Jones databases. Currently, hotlinks in the Wall Street Journal Interactive Edition articles generally take you to Dow Jones' own Business Briefing reports for more background information on companies -- material included in the $49 annual fee. Sometimes it refers to an article in the Publications Library for which Dow Jones does charge extra. If a reporter links to an outside Web site, the WSJ Interactive Edition will put the URL in parentheses in the text of the article.

In contrast, IAC's InSite Pro Web-based service, introduced in 1997, offers tremendous assistance to CI research through its use of hyperlinks. Top IAC databases (including PROMT, Trade & Industry, Newsletter, Computer, and Company Intelligence databases) are searchable in easy-to-use, simple, or advanced interfaces. But the features of greatest value appear in the resulting full-text articles. Running along the left side of the screen, IAC has included hyperlinks to other stories gathered and categorized by IAC's vast index fields. For example, one article retrieved in a search on Gillette and antiperspirants displayed hyperlinks to stories about the named advertising agency, as well as to the antiperspirant and advertising industries. In addition, when available, IAC includes hyperlinks to corporate Web sites in the Company Intelligence Database (currently 9,000 links). The company plans to include corporate hyperlinks in the remaining databases in 1998. By harnessing the Internet's innate hyperlink function, InSite Pro makes intelligent connections, conveniently extracting and offering many of the resources we, as CI experts, would pursue in additional new searches. With its flat-fee pricing, InSite Pro allows you to follow those hyperlinks to your

heart's content.

Manning & Napier (http://www. mnis.net) is another new entrant with a flat-fee, Internet-based service called DR-Link. The service offers several useful features, particularly its natural language search capabilities and assistance in focusing a search. In 1997, the service introduced a new "Visualizer" Java function that can prove to be a great asset in CI research. In response to a query about biotechnology research in vaccines, for example, the system produced a hyperlinked chart, visually displaying and separating out companies and organizations, people, countries, and subject areas related to the search -- all valuable facets for a typical CI project.

Northern Light (http://www.northern light.com) represents one of those fresh new Internet faces mentioned above, in this case, a company with a novel idea for a commercial online service -- a highly accurate yet simple search engine offering inexpensive full-text articles from more than 400 business and industry publications combined with targeted Web searching. Based on a mass marketing model (sales in the millions of transactions to the general public as well as professional researchers), Northern Light's pricing structure charges per article, with no sign-up or online searching fees. Northern Light also offers features that ease the CI research process. Search results, automatically ranked for relevancy, include a list of custom search folders which hyperlink to stories categorized by subject, industry, publication, etc. Again, these hyperlinks help focus quickly on results, eliminating the need for new searches. Northern Light's excellent search engine can also search the Web with highly accurate results.

A new Web site called CEO Express (http://www.ceoexpress.com) is a business researcher's dream. Completely revamped this year for more intuitive use, the site features one-stop shopping with a vast, well-organized collection of essential company- and business-related sites and services on a single page. Aimed at "busy executives," CEO Express not only links to the more obvious news, government, and business sites, but also connects to overlooked sites such as Chambers of Commerce, the Internet Bankruptcy Library, and software download sites.

Last year, the online industry news story with the greatest coverage was clearly M.A.I.D. PLC's purchase of Dialog from Knight-Ridder. It's too soon for the impact from the sale to reach the searcher. Latest reports promise a marriage of Profound's intuitive Web-based interface with Dialog's vast databases in the future. However, in 1997, the big push was DialogWeb. This product offers little of inherent value to information professionals performing CI research.

However, Dialog's addition of company index fields to close to 80 files (including IAC Newsletter, Early Edition, McGraw-Hill, and American Banker, among others) has proven to be a great help, allowing index searching for company names in files whose producers did not see fit to include this basic descriptor field. Dialog-

generated company names have now gone onto almost every news source added since January 1997.

Dow Jones Interactive, whose Publications Library at presstime extended to over 5,000 titles and 70 million articles, also instituted a vendor-based company indexing system built on sophisticated software. The indexing provides standard company name and ticker symbol access across the whole range of news sources, regardless of whether the database producers provide their own company name descriptors.

In a "marriage of convenience," Dow Jones, the Financial Times, and Dialog allied to produce the World Reporter. When the World Reporter database reaches its full strength (currently it includes over 400 publications), it should prove worthwhile for international CI research, offering a company index field, news from 500 major international publications in 26 languages (currently 17); and translated abstracts or full-text articles in English.

The Net's Database Evolution

With most major companies supporting home pages, the Internet has grown into a vital information resource unto itself -- not just an aggregator or alternative platform for commercially available databases. The Internet now serves as a data source itself, combined with a searchable advertising medium. Although companies clearly hype their products, image, stock, etc., they indeed provide valid and valued information through their Web sites. Useful CI information such as news releases, SEC filings, job postings, executive biographies, conference participation, speeches, all commonly appear on company home pages. Moreover, even "hype" can have value -- knowing how your competitor portrays itself, its products and services is fundamental to understanding your competition.

At Fuld, we typically begin research projects with a thorough review of a company's home page. You can often guess the home page for a company (http://www.companyname.com). Garden variety search engines often have difficulty zeroing in on a large company or a firm with a common name (imagine looking up Apple's home page). So when the shortcut try doesn't work, we use several company look-up Web sites. Fortunately several new services have arisen to replace the Domain Name Lookup Service, which managed to redesign itself into inadequacy last year (http://www.internet.org). Several new favorites include NetPartners Company Locator, which uses Internic registration information, thus offering the most extensive database (http://www.netpart.com/resource/ search.html). Several other sites focus on larger, predominantly public companies: Hoover's (http://www.hoovers.com); CompanyLink (http://www.company link.com); Companies Online (http://www.companiesonline.com); and the multi-function, ubiquitous Yahoo! (http://www.yahoo.com/Business/Companies).

Drilling Down Locally

Local press, always a fertile field of useful information, offers intensive and extensive coverage of hometown companies. Revealing stories on company activities as major as reorganizations or as minor as OSHA inspections will receive more coverage in the local press than elsewhere. Dow Jones continues to be the leader in this arena, expanding full-text coverage to more than 500 newspapers, both large and small. However, the Web-based Dow Jones Interactive service eliminated a useful feature from its news retrieval system -- the ability to scroll through a list and limit your search to a whole state or a single newspaper. Instead, with the Web-based service, you must know and use obscure keyboard commands (such as sn=boston globe) to search individual newspapers. However, according to Dow Jones, the company plans to bring this functionality back to the Web product next year. On the other hand, only the Dow Jones Interactive product features relevancy ranking and a free peek at the first sentence or so for each article.

Dialog also continued its increase of local press coverage to more than 100 local papers. Weak spots such as Texas, which had lacked current news in the Dialog collection, now have local press coverage.

Sadly, 1997 saw the passing of a local press pioneer. DataTimes, an early advocate of local press offerings, ceased separate operation and was integrated into UMI's Proquest Direct product. Under an agreement with Dow Jones, UMI will direct DataTimes customers to Dow Jones Interactive, which will pick up local press formerly offered by DataTimes.

LEXIS-NEXIS, the pioneer of full-text news online, added Bloomberg News in 1997 with corporate news coverage from around the world, along with extensive reporting on governments and financial markets. They have over 60 of the nation's top 100 newspapers available full text, with 15 local papers added just in 1997, along with lots of wires and business news sources. LEXIS-NEXIS appears to have a steady growth policy when it comes to U.S. and international news sources.

Hometown papers also have a strong showing on the Internet. For one client, we sought information about a pilot energy project in several small New Hampshire towns. Precious little appeared on the conventional online services; even Dow Jones at that time had few full-text offerings from New Hampshire. Instead, we turned to the Internet and ultimately found crucial information, including invaluable contact names, through small local newspapers on the Web.

Newsworks (http://www.newsworks. com), the first site offered by New Century Network, is a joint venture of nine publishers featuring 125 newspapers as well known as The New York Times and as obscure as the Anchorage Daily News. The beauty of the site is that

you can search all papers through one interface with one syntax and retrieve free full-text articles. The beast of the site is that you still never know if you'll retrieve a full-text article or an outdated dead link to a non-archived article.

If your situation requires that you know full-text costs beforehand, you can head over to Newspaper Archives on the Web sponsored by the Special Libraries Association News Division (http://sun site.unc.edu/slanews/internet/archives.ht ml). This site features a chart with per article charges plus hyperlinks to full-text archived newspapers. For those with a subscription to BiblioData's resource-filled CyberSkeptic's Guide to Internet Research or Fulltext Sources Online, you also have access to BiblioData's "Private Zone" (http://www.bibliodata.com/private.html), a password-protected collection of more than 400 newspaper and magazine sites with archives both free and full-text.

American City Business Journals (http://www.amcity.com), a site noted for local and regional business news, features 35 regional business papers. A major revision this year increased its utility. This is one of the few sites offering free full-text archives and a single searchable interface. Sometimes the Internet still offers a free full-text cake and lets you eat it with a single searchable interface!

Clipping Without Scissors

Electronic alerting and clipping services can also prove very useful to CI researchers. It's hard to argue with an e-mail alert about a competitor waiting in your mailbox when you arrive at work in the morning. This year has seen several changes among Internet-based alerting services. The most promising change is the extensive service upgrade of Farcast, which also renamed itself INQUISIT (http://www.inquisit.com). Once limited to wire services, INQUISIT now serves up inexpensive Internet-based alerts and industry "broadcasts" derived from 400 newspapers and periodicals around the world. Simplified templates help set up a personalized alert on a company, product, or name. Headlines, paragraphs, or full text appear directly in your e-mail box or are retrievable via the Web. The INQUISIT broadcast service delivers a list of industry-specific headlines several times a day, allowing you to check off the stories you want, which they then deliver to your mailbox, full text, within 20 minutes.

The creation of NewsEDGE (http://www.newsedge.com), formed by Desktop Data's merger (dare we say acquisition?) with Individual, Inc., prompts many questions. While Desktop Data's high-priced offerings focused on real-time, corporate-wide news delivery, Individual's claim-to-fame centered on inexpensive or even free personalized news. The newborn NewsEDGE will have more than 2,000 content sources under license. The merger may be good for the companies involved, but the endurance of Individual's affordable services remains to be seen. To quo Han Solo, "I don't have a good feeling about this." After all, how will we, as customers, benefit from

the high-priced spread merging with its low-cost competitor?

A new wrinkle in alerting services can help researchers monitor competitor Web sites. Offline Web browsers that download full pages from a Web site are predominantly designed to peruse Web pages offline. However, some will also send alerts when specified Web pages (including keywords or passages) change. You can set up some offline browsers to monitor a competitor's Web site for changes on pages holding SEC filings, news stories, job postings, new products, management changes, executive speeches -- all keys to competitors' strategies. Two commercial packages of note are Tierra Highlights2 from Tierra Communications (http://www.tierra.com) and Web-Whacker from the Forefront Group (http://www.ffg.com). Both software publishers upgraded their products last year. For lists of other offline browsers, visit Tucows (http://www.tucows.com/us. html), a well-regarded gathering place of Internet software. When prompted, identify the continent, country, state of your location, and operating system. Then choose Offline Browsers from the chart of available software.

Patently Clear

Patent searching is critical to an effective CI research program. Patents point to new areas of R&D, to products slated for future commercialization, and to inventors/experts useful as contacts. Several Internet-based innovations now make patent information more accessible to a wider audience. Introduced early in 1997, IBM's Patent Server (http:// patent.womplex.ibm.com) offers several advantages over the USPTO patent site from which it derives its information, particularly ease-of-use and free full-patent images. The site directly links to Optipat, which sells full-text patents for as little as $2.50 each.

MicroPatent and its Internet-based patent site (http://www.micropat.com) offer other advantages. MicroPatent takes a step forward by offering a fee-based e-mail alerting service based on patent applications filed with the European Patent Office (EPO), World Intellectual Property Organization (WIPO), and Japanese Patent Office (JPO). This service can offer a competitor intelligence advantage: U.S. patent information is only released when a patent is granted -- often several years after the filing of the application. EPO, WIPO, and JPO, on the other hand, release information within 18 months of the application filing -- sometimes providing a competitive leg up, so to speak.

Small Is Beautiful

While we often expect major new services to provide the greatest utility, sometimes the small things offer the greatest value. For many of our research projects at Fuld, we routinely retrieve SEC filings, seeking invaluable information concerning financials, operations, products sold, markets, management, and much more. Competitors' SEC filings are fundamental competitor intelligence

research tools.

Several companies now offer inexpensive or free SEC- and Internet-based alerting services. Who Where? (http://www.whowhere. com) sponsors EDGARAlert! (http://www.whowhere. com/EDGAR), a free service that sends you e-mail when one of your specified companies submits a new SEC filing, though there may be a time delay. In contrast, EDGAR Online's Watchlist (http://www.edgar-online.com/) service is a low-cost, real-time alerting service. As soon as the SEC filing is submitted, you receive an HTML-enabled alert in your mailbox. Click on the hyperlink and you go directly to the filing in question. It also features a new service called EDGAR Online People that allows customers to search SEC filings by an executive's name and thus uncover position held within the corporation, compensation packages, and board memberships.

The Help-Wanted Intelligence Explosion

While numerous Internet services abound offering inexpensive alternatives to the "high-priced spreads" of commercial online services, the Internet does offer new and unique services. For example, job postings provide a wealth of clues to a company's current technology use, research interests, and future plans. Identify the persons and skills your competitors hire today and you will often uncover their future. In years past, this information was only available through specialty database companies, newspaper clipping services, or by subscribing to newspapers in competitors' home towns. Today, a myriad of ways can help identify hiring trends, including competitors' home pages and searching Usenet group posting via Deja News (http://www.dejanews.com).

Commercial job sites such as Monsterboard (http://www.monsterboard. com) and CareerMosaic (http://www. careermosaic.com) can be worth their weight in gold. Companies frequently post lengthy lists of various and varying jobs, offering a peek at the range of their work. Monsterboard now offers an alerting service aimed at job seekers that can also help the CI researcher. You can specify company name or type of job and receive an e-mail alert when that company posts a job.

For smaller companies, or those not yet Internet-savvy, CareerPath (http:// career.careerpath.com) is a unique resource which expanded coverage last year. Several years ago, CareerPath pioneered Internet job postings by offering two weeks of job advertisements from six major newspapers across the country, searchable via one interface. The site still allows a two-week search by company name or job category, but now includes more than 40 large and small newspapers across the country and imports job postings from major corporate Web sites.

Many people err in believing that Japan is our largest trading partner, when in fact that honor belongs to Canada. Thus, a welcome addition to business research on the Internet was last

year's start-up of Canada's corporate filing site, similar to the U.S. EDGAR site. Called SEDAR, (System for Electronic Document Analysis and Retrieval), the bilingual service is free and includes most filings required of Canadian public companies going back to January 1, 1997 (http://www.sedar.com/).

Last year also saw the death of another valuable commercial service known for unusual and sometimes unique information. NewsNet, noted for its collection of hard-to-find industry newsletters, including some newsletters unique to its service, ceased operation in August 1997. Its remaining Web site now prompts former subscribers to a Web-based online service called BrainWave, offered by WinStar Telebase. This service features an interface that mimics NewsNet's old Baton interface. However, BrainWave's newsletter offerings are limited to those included in IAC databases. Due to NewsNet's demise, several technical newsletters are no longer available electronically -- commercially or via the Internet -- a tremendous loss for researchers in those industries.

And Now 1998...

As to the future for competitor intelligence and online resources? Remember that line from The Graduate: "Plastics"? Just do a global replace with the word "Internet." A clear message has been sent: Proprietary systems will see few new developments. Look at the highly visible migration to the Web with DialogWeb and Dow Jones Interactive. The Internet is the clear winner of the platform race.

When used to its full potential, the Internet offers competitive advantages over older systems. Expect traditional market forces to apply, however. New entrants must offer clear advantages. In the Internet world, creativity is far easier to accomplish when you are a brash upstart, not hamstrung by commitments to older technologies. Look to the rebels for sizzling innovative products offering such desirable features as intuitive interfaces, access to new resources, unique services, time savings, lower costs, or a combination of the above. We will all adjust to the thrills, challenges, and exasperation inherent in this brave new Internet world.

America's Network Sept 1, 1998

Mixed nuts: automated provisioning, HFC and IP. (hybrid fiber-coax network; Internet protocol)

Author/s: Throop Wilder

Building a self-service subscription system for cable modem Internet services.

Customer self-provisioning is not just an idea, it is a reality. However, it is a reality fraught with complexities, especially when it comes to cable modem-based Internet services. This has been the experience of MediaOne (Englewood, Colo.) as it has encountered the challenges of automating subscriber sign-on, service selection and dynamic provisioning of its national, cable modem-based Internet service.

MediaOne has worked closely with American Internet Corp. (Bedford, Mass.) to deploy a carrier-class Internet protocol (IP) service infrastructure, and has recently formed a venture with Time Warner Cable's (Stamford, Conn.) Road Runner division to jointly deliver data service to both companies' subscriber base - roughly 100,000 Internet users out of a total of 15 million subscribers.

In planning for the service's requirements, growth projections for data-over-cable services indicate that cable operators must look to highly automated systems. What are the facilities offered by IP networks that enable or impede automation? How much subscriber-to-customer service representative (CSR) interaction is desired? When can the installer truck rolls stop? These are the types of questions being asked by all operators as they expand into data services over the cable network.

But what are the goals of an automated provisioning system?

Lower administrative costs - As we will discuss below, signing up a subscriber today is still a very expensive proposition. In fact, the payback time for many multiple systems operators (MSOs) ranges in the six-month to one-year time frame. Even once the signup cost

has been recouped, a help desk call into a service center may cost between $40 and $75, and wipe out one to two months' revenue for that specific subscriber. Most costly of all is the series of calls typically required to fix a user's billing problem.

Therefore, a subscriber-provisioning system that can automatically address these "life cycle" events can literally make the difference between the red and the black.

Provide flexibility for market trials and new service rollouts - Current systems are not geared to the types of marketing campaigns MSOs want to run. For example, in an automated system, the MSO could easily test multiple free-trial scenarios by pointing subscribers at different signup Web pages, and get almost instant feedback on the trial's efficacy.

Another (possibly more serious) constraint with current systems is that it takes most MSOs about a year to roll-out a new service. This practice will not scale as subscriber growth rates continue to skyrocket. Cable operators expect IP telephony to take off within the next couple of years, but, with current provisioning systems, the backlogs could stretch into tens of thousands of subscribers.

Improve the customer experience - One of the difficulties in signing up for cable service today is scheduling the installation of service at the home, especially with the number of two-working-parent homes. However, give subscribers flexibility in scheduling by choosing from a range of options through a Web interface at their own rate, and new subscribers can feel as if they have greater control over the signup experience.

SPECIFIC CHALLENGES OF BROADBAND IP NETWORKS

Reaching these goals is no easy feat. Anyone in a major metropolitan area today either has, wants or knows someone else with Internet-over-cable service. The blazing speeds of 1.5 Mbps render the Web experience what it truly should be - serving up graphically intensive pages in the snap of a finger. Such services can change the World Wide Wait into the World Wide Whiz.

Furthermore, the predicted domination of high-speed Internet service by the telcos over the cable industry has not happened. Other events, such as AT&T Corp.'s (New York) purchase of TeleCommunications, Inc. (TCI; Englewood, Colo.), further underscore the strategic value of the cable industry in reaching consumers with new data services. Cable modem service's high-speed, reasonable monthly rates (around $40) have resulted in subscribers signing up at the rate of thousands a week. Unfortunately, actually getting the service isn't so snappy.

THE SIGN-UP PROCESS TODAY

All cable operators still rely on administratively intensive processes for signing up and activating a subscriber. Typically, the cycle can take up to several weeks, and follows a path something like this:

1. The potential subscriber calls CSR.

2. The CSR takes the subscriber's information and initiates a work order for installation. The CSR and subscriber arrange an installation date.

3. Two field technicians - one skilled in cable networks, the other in home computers - drive to the subscriber's house.

4. The cable technician performs several tasks depending on the current state of the cable plant in the subscriber's neighborhood. In most cases, the cable technician replaces an existing high-pass filter (to block upstream noise from household appliances) with a combination filter and two-way splitter. The splitter connects one coax cable to the subscriber's set-top box and the other to the cable modem. In step two, the technician then installs the cable modem. An additional wiring step may be necessary if the customer premesis wiring can't support data demands.

5. The cable technician then checks the network hardware address of the cable modem and calls to relate this information to the CSR, who enters the information into a database of cable modems that are authorized to connect to the hybrid fiber-coax (HFC) network.

6. The computer technician checks the user's PC to make sure it supports IP has an ethernet card, and installs the client software. In addition, the PC technician must configure the software to make sure the settings are correct. Working with the cable technician, the PC technician then connects the PC to the cable modem.

7. If all goes according to plan, the subscriber is now "live" on the network.

Although MediaOne and other MSOs don't make public their provisioning costs, the steps outlined above clearly indicate high signup costs which require extended time to recapture.
Also, signup is not the only high-cost subscriber/CSR interaction. Service changes, upgrades, moves, billing lapses and other subscriber lifecycle events require human intervention, and hurt the operator's bottom line. All this in a business which traditionally operates within relatively thin profit margins and great attention to cash flow.

A NEW MODEL FOR SUBSCRIBER PROVISIONING - THE SERVICE ZONE

Fortunately, there is good news. An explosion in new technologies

and standards is making possible a truly automated self provisioning system - one in which users perform many registration and provisioning steps themselves.

The key to the solution is the development of a new concept called a "service zone" ([ILLUSTRATION FOR FIGURE 1 OMITTED], page 30). Traditionally, when a user signs up for service, it is either on or off. It is on if you pay, off if you don't. This concept is simple, but frustrating, because switching users on or off requires them to interact with CSRs.

The service zone creates a new state where service is neither on nor off. Instead, the subscriber has just enough "on," or network access, to solve their own problem. Ideally, a subscriber should be able to buy a cable modem at a retail outlet, set it up at home, and immediately gain access to the network through a Web browser. At this point, they would not be able to surf the 'Net, but they would have enough access to sign up for the service. Once authenticated and proven credit worthy, the subscriber's PC and cable modem would be re-provisioned to grant the subscriber full access to the services for which he or she subscribed. At this point, the subscriber has exited the service zone.

Imagine now that the same subscriber has failed to pay the bill. In the automated provisioning model, the subscriber is dynamically placed back into the service zone. At this point, the only resource available to him or her is a billing server, which greets the user with a message such as, "Hello, your account is now 'X' days past due. Please type in a valid credit card number, and your full access will be re-enabled."

This concept isn't new. Interestingly, users of America Online Inc's. (AOL; Dulles, Va.) services already have this capability available to them today. However, AOL controls both the client and server software, and its proprietary software provides the back-office link to establish and update service.

However, in the open, Internet space, service providers do not have the luxury of controlling the end-to-end technology. The only safe assumption about client software is that a user has an IP-capable system and a Web browser.

Through the development and elaboration of a service zone in the cable industry, many CSR interactions with subscribers can be completely eliminated, while field technician time at the subscriber residence can be dramatically reduced. Additionally, a great deal of plain-old human error is removed from the system. The result, ideally, is a painless subscriber sign-on experience, and a smoothly functioning, low-cost administrative model for the cable operator.

SERVICE ZONE BUILDING BLOCKS

The technology foundation for automated provisioning in broadband

networks like MediaOne is made up of elements which are, above all, scalable and flexible. On the network side, the service depends on fast-changing parameters, such as core IP addresses and domain names. Imagine if your phone number changed twice during one conversation - the phone company systems would have to know who had what number, when, and for how long. This is the kind of feature required in dynamic IP networks.

On the service-management side, systems have to interoperate through well-defined interfaces. This means that a customer-care application must be able to talk to a network-provisioning server, which must be able to output key data to a variety of systems, from billing to network management.

Although MediaOne still has much integration work ahead, it already has many of the important building blocks in place. These key components include:

Web-based subscriber interface - MediaOne has developed Web interfaces for administrative personnel who must make rapid, and sometimes complex, changes to the provisioning infrastructure. For example, as subscriber growth rates increase, networks frequently need to be renumbered. MediaOne's network managers use Web tools to enter changes to the system. The Web system then interacts with other services to orchestrate the simultaneous changing of network numbers across multiple services.

It is this same concept that MediaOne seeks to introduce to the subscriber experience. Essentially, Web interfaces, at their best, represent role-based, workflow systems where each screen of information shows users only what they needs to know (role-based), and leads them, step-by-step, along the road of options appropriate to the task they are performing (work-flow).

For the MSO, this means new customers get one screen, while subscribers with lapsed bills get another. In an automated process, the system tells the Web interface which screen to present. In cases where a field technician acts as a "proxy" for subscribers, the screen can be tailored dynamically to lead them through the appropriate actions for their services. This becomes extremely important as data over cable service interface specification (DOCSIS) modems begin hitting retail outlets and service technicians from those stores perform some of the initial PC setup.

IP address provisioning - To provision name and IP address, MediaOne is using Network Registrar ([ILLUSTRATION FOR FIGURE 2 OMITTED], page 32). These name and address attributes uniquely identify devices and users on a network. For example, when a cable modem is installed, both it and the subscriber's computer are automatically provisioned with IP addresses that make them reachable on the network. In addition, many other devices on the data network have to know that the subscriber device addresses are

valid and authorized for the particular network. To solve these security concerns, the system has several features which relate a subscriber's PC to the cable modem and prevent "spoofing," in which unscrupulous users steal service from paying users.
In addition to simple address allocation, the system also can distinguish among different classes of service and place subscriber devices into what are sometimes called "policy groups" - groups of devices within a particular address range that share a certain network bandwidth policy.

Network Registrar propagates the address information throughout a network, because that data must be known by all network devices and services. The system also updates both lightweight directory access protocol (LDAP)-based directories and domain name service (DNS) servers with real-time address and name information.

Where service-management is concerned, in a service provider environment, the "user" of a system is as likely to be another program as it is a human. In an automated provisioning environment, this is even more true; once a subscriber signs up for service over the Web, that same Web registration system must communicate with address provisioning system to make sure that the user gets a valid name and address which map to their service level.

Performance goals - The value of specifying performance goals for each system requirement should not be ignored. For example, how many seconds should it take for a subscriber to get added to the central data store? What network load is generated when an entire neighborhood is hit with a power failure, and 1,000 network devices (computers and modems) try to reboot simultaneously? Using various mathematical models, MediaOne and American Internet have mapped traffic distribution patterns and tied them back to subscriber growth patterns. The understanding of these patterns in turn will allow MediaOne to predict deployment requirements for new software and hardware components in the infrastructure.

NEXT STEPS

Within the last three months, a cross-functional team representing network and customer service functions within MediaOne has begun meeting to work with the network elements in place and design an architecture which will scale to meet both the service requirements while continuing to handle the high subscriber growth rates.

Throop Wilder is founder of American Internet Corp. (Bedford, Mass.). Stephen Van Beaver is vice president of customer care and operations at Media One (Englewood, Colo.). Comments regarding this article should be addressed to anrespond@americasnetwork.com.

AI Magazine Summer, 1998

Mobile digital assistants for community support.

Author/s: Yoshiyasu Nishibe

* We applied mobile computing to community support and explored mobile computing with a large number of terminals. This article reports on the Second International Conference on Multiagent Systems (ICMAS'96) Mobile Assistant Project that was conducted at an actual international conference for multiagent systems using 100 personal digital assistants (PDAs) and cellular telephones. We supported three types of service: (1) communication services such as e-mail and net news; (2) information services such as conference, personal, and tourist information; and (3) community support services such as forum and meeting arrangements. After the conference, we analyzed a large amount of log data and obtained the following results: It appears that people continuously used PDAs in their hotel rooms after dinner, e-mail services were used independently of the conference structure, but the load on information services reflected the schedule of the conference. Postquestionnaire data showed that our trial was considered interesting, although people were not fully satisfied with the PDAs and services provided. Participants showed a deep interest in mobile computing for community support.

The project reported in this article is intended to support, using mobile computing, increasingly diverse and amorphous groups of people. For this purpose, mobile computing services should change from applications based on point-to-point communication to those based on communitywide communication. Community mobile computing is feasible because the number of users of wireless communication services is rapidly increasing because of the falling prices of mobile terminals and cellular telephones. We show that mobile computing technology supports international conferences effectively and combines well with conventional desktop computing technology.

Internet access services are now popular at international conferences. Several tens of desktop terminals are usually provided

at a conference site. However, the services are constrained in both space and time; the services are only accessible from terminal rooms during the daytime. However, the salient feature of mobile computing is that users can use their personal digital assistants (PDAs) anytime and anywhere. The challenge of this article is to apply mobile computing to an actual international conference and investigate its role quantitatively.

Previous work in computer-mediated communication was classified into two main categories (Mynatt et al. 1997): First is multiuser dungeons (MUDs) and casual meeting tools that support a virtual space where everybody can meet and talk. The other is MEDIASPACE (Bly, Harrison, and Irwin 1993), which is a hybrid of virtual and real spaces. Mobile computing, however, aims to amplify interactions in a real space and should therefore be classified in a third category.

Although internet access services are common at international conferences, no empirical study on them has been reported. Similarly, in the mobile computing literature, several papers have been published concerning technologies and applications, but no analysis on social interactions has been published (Weiser 1993). Human-computer interaction (HCI) researchers have expended substantial effort on multimedia communications and conferences for relatively small teams (Olson, Olson, and Meader 1995). Several reports address large groups in business applications (Sproull and Kiesler 1988), education (O'Day, Bobrow, and Shirley 1996), and home computing (Kraut et al. 1996). However, there is no report on mobile computing with a large number of users.
In this article, we first report the experiment called the Second International Conference on Multiagent Systems (ICMAS'96) Mobile Assistant Project. This conference was held in Kyoto, Japan, from 9 to 13 December 1996. The project provided (1) e-mail and internet access services; (2) conference, personal, and tourist information services; and (3) forum and meeting arrangement services. In tins experiment, about 100 personal digital assistants (PDAs) with wireless telephones were loaned to conference attendees without any charge to try out the system. To the best of our knowledge, this was the world's first application of mobile computing to community support.

By reducing the time and space constraints of desktop computing, mobile computing guarantees freedom of use of network services. As a result, the demand for various information services was clearly revealed. We show that mobile-computing technology supports international conferences effectively and combines well with conventional desktop computing technology.

ICMAS'96 Mobile Assistant Project

The goal of the ICMAS'96 Mobile Assistant Project(1) was not only to support communication services such as e-mail but also to provide,

through PDAs, a variety of information required at international conferences (figure 1). PDAs were provided to about 100 people, roughly one-third the conference attendees. People could use them in the conference site, lobbies, and hotel rooms as well as outdoors.

[Figure 1 ILLUSTRATION OMITTED]

Figure 2 describes an overview of the system configuration. The server machine was connected to the internet to provide e-mail exchange and information retrieval. The server system was programmed with TELESCRIPT, which is a language intended for agent-oriented mobile computing applications (White 1994). The client system was built on MAGICCAP OS running on MAGICLINK.(2) The server system was programmed with TELESCRIPT (White 1994),(2) a language intended for mobile computing applications. The client system was built on MAGICCAP OS running on MAGICLINK. Each client could connect with the server system used by wireless public telephone lines. The data-transmission rate of wireless lines on the physical layer was 9600 bits a second (bits/s).

[Figure 2 ILLUSTRATION OMITTED]

The project became larger than we first expected. Thirty telephone lines were newly introduced and connected with the server machine placed at the conference site. We then found a serious communication problem between the PDAs and the server. Because mobile communication traffic in this experiment differed from typical telephone services, we found that the existing facilities were not enough to cover the demand. After negotiation with the local cellular telephone company, the situation was substantially improved by allowing several tens of calls from the conference site to be conducted simultaneously.

The project was announced to conference attendees beforehand so that they could apply for a PDA. We asked all participants to fin out a prequestionnaire together with the application form. At the conference registration desk, after the process of authentication and contract, the PDAs and the telephones were handed to the project participants.
Figure 3a shows a participant receiving a client system at the registration desk. The contract included permission to use log data for academic purposes. During the conference, a help desk was set up at the conference site to assist participants with various technical problems. Figure 3b depicts a participant using the service while he stands in the lobby of the conference hall. Figure 3c also depicts a participant using the service in the event hall where the conference was held. Figure 3d depicts the help desk. Figure 3e is a photograph taken during the excursion to Nara Historical Park, which was done as a part of the conference schedule. Help-desk staff also traveled to Nara Park during the excursion. After the conference, all PDAs and cellular telephones were returned without trouble.

[Figure 3 ILLUSTRATION OMITTED]

To design the mobile digital assistant as an agent for community support, we first classified an individual's activities into the following three categories: (1) communicating with other people to share information, (2) interacting with the real world to determine his/her actions, and (3) understanding community activities to identify his/her role in the community. We then designed various functions to support these three types of activity and implemented (1) communication services, (2) information services, and (3) community services, as described in table 1. These services consisted of commercial services such as e-mail and various original services developed by about 20 engineers and students from NTT Information and Communication Systems Laboratories, Kyoto University, and the Nara Institute of Science and Technology The original services given in table 1 were provided during the conference:

Table 1. Services Provided.

Communication services		
	Send message	Messages are sent from PDAs to other participants or internet users through telephone lines.
	Receive message	Messages are received by PDAs from other participants and

internet users

through telephone

lines.

Information services

Conference

Information Session schedules

and maps are

stored in each

PDA. Abstracts of

presentations are

provided through

telephone lines.

Personal

information COMMUNITY VIEWER

provides

participants'

personal

information stored

in each PDA and

visualizes the

interactions among

the participants

through telephone

-line transfer.

Tourist information

ACTION NAVIGATOR

provides

information

(stored in each

PDA) such as shops

and sightseeing

spots.

Community services

Forum and meeting

INFOCOMMON helps

users to exchange

knowledge and

ideas through

shared card

information

through telephone

lines. SOCIAL

MATCHMATER

provides a way to

find other people

who have similar

interests.

Statistics feedback

The current state

of service use is

shown in real

time through

telephone lines.

PDA = personal digital assistant.

First, ACTION NAVIGATOR is a system that supplies tourist information and additional information to help the user make decisions.

Second, INFOCOMMON provides a weak information structure for human information sharing. The weak information structure connects various information media without defining the semantics rigorously.

Third, COMMUNITY VIEWER visualizes interactions among participants and provides a better view of community activities. Figure 4 shows the screen image of the opening index for selecting these services. Because most services were developed just for this experiment, we were able to embed enough codes to obtain log data such as the log for access to the server and changes in PDA output. We also asked all users to fill out a prequestionnaire and a postquestionnaire to find out about their usual computer environments and impressions of the project and the services provided.

[Figure 4 ILLUSTRATION OMITTED]

ACTION NAVIGATOR: Supporting Action Making

We are always making decisions, such as selecting the restaurant when we eat out (Engel, Blackwell, and Miniard 1990). Typical

tourist information services provide objective information such as name, telephone number, address, and business information when users request information.

However, depending on search conditions, they often return a large number of search results. In this case, the user has to choose the one he/she wants from this large amount of information. It is difficult to choose on the basis of objective information only. In general, when we make decisions, we often refer not only to objective information but also to subjective information, for example, recommendations of others or reviews in newspapers or magazines.

ACTION NAVIGATOR (Ohtsubo et al. 1998) supports decision making with subjective information from others coupled with a tourist information service. Figure 5 shows an example screen Of ACTION NAVIGATOR. Information spots are displayed in the map of Nara city. Users could get detailed information on shops, restaurants, and so on, by clicking on the corresponding spot

[Figure 5 ILLUSTRATION OMITTED]

The subjective information in ACTION NAVIGATOR was based on answers to the following two questions: (1) How often do users look at the detailed information for each information spot? (2) Do users post news about this information spot? Each PDA counts the number of times that the user refers to the detailed information and sends this count to the server. The server gathers the number of references and calculates an evaluation value (we call it the active level) for each information spot.

When starting Up ACTION NAVIGATOR, users are asked whether or not the system should update the active levels. When the user chooses to update these levels, the PDA communicates; with the server to get the current active levels and changes the size of each information spot according to these levels. The active level took four values in this experiment. The information spot that has the highest active level (displayed as the biggest icon) is called the hot-spot. We examined the accesses to the information spots. It appears that 50 percent of all accesses (furthermore, 80 percent of the first accesses) were to the top 7 percent of the hotspots. The results show that the simple hotspot mechanism effectively realizes social filtering of tourist information.

However, users did some decision making during the conference independent of using ACTION NAVIGATOR. We asked users the following question in the postquestionnaire: 'if you went to some restaurants or sightseeing, why did you select them?" Figure 6 shows the answers to the question. According to this chart, users' decisions depended on the recommendations of others: Individual preference of others takes an important role in making decisions.

The result supports our policy of designing ACTION NAVIGATOR.

[Figure 6 ILLUSTRATION OMITTED]

INFOCOMMON: Sharing Information among Participants

The effective use of tacit knowledge or nonlinguistic means of communication would be an interesting issue in addition to unambiguous representation of information. It is not until recently that these issues have been addressed from computational points of view.

The basic assumption behind INFOCOMMON (Maeda et al. 1997) is that information sharing facilitates human interaction and group formation (Gaines and Shaw 1994). Talking about a common hobby might result in an academic discussion. To facilitate information sharing among participants, we used several design principles: First, we tried not to enforce any one particular concept. Instead, we allowed a lot of freedom in the use of terms and the structure of shared information and freedom to incorporate useful information from various viewpoints. Second, we made the information space a single seamless space, which releases the user from working with a rigid menu. Third, we enabled the user to build a personal information space where he/she can organize relatively small amounts of information, as desired (Nishida et al. 1995). This approach, we believe, makes shared information visible and encourages human interaction and community formation.

Technically, we propose a weak information structure for encoding shared information The weak information structure connects a wide variety of information media, such as natural language text, hypertext, and images, without defining the semantics rigorously. By leaving ambiguity and vagueness in the representation, we can incorporate a wider variety of useful information into the information base, ranging from formal conference information to articles that might be useful in finding interesting restaurants or points of interest.

The information base Of INFOCOMMON is a collection of information cards organized on this weak information structure. Representative keywords are automatically extracted from each information card. The collection of information cards is organized by shared keywords and predefined relations. Given a set of keywords, INFOCOMMON responds with the set of information cards connected to the keywords and provides a visual interface for retrieving and sending information cards. Figure 7 shows an example screen Of INFOCOMMON. The relation between two information cards is displayed by a link. The result of retrieval is stored in the user's local information base where the user can rearrange the collection of information cards and add and remove nodes and links, as desired.
[Figure 7 ILLUSTRATION OMITTED]

We analyzed how INFOCOMMON was used by log files and questionnaires. Figure 8a shows the answers to the question, 'For what did you used INFOCOMMON?" Fifty-nine percent used the system for information retrieval and 19 percent for information sending; sending comprises news, help desk, and discussion. The reasons INFOCOMMON was used for news "because keyword search was easy and useful" (14 persons), "because I found interesting topics in INFOCOMMON" (5 persons), and "because I had a question" (4 persons). We determined that INFOCOMMON added a new facility to the conventional news reader. Major reasons for the choice "didn't use" were "slow information retrieval" (8 persons) and "I couldn't connect to the server" (4 persons). These problems involve server response, so they can be corrected easily.

[Figure 8 ILLUSTRATION OMITTED]

Figure 8b displays answers to the question, "For what topic did you use INFOCOMMON?" About search results, 81 percent felt that the search results were fine, and 51 percent answered that they were satisfied with INFOCOMMON. About the usefulness of INFOCOMMON for information retrieval and discussion, 55 percent answered that INFOCOMMON was useful for getting the information they needed. However, only 26 percent answered that it was useful for discussion. We feel that the five-day period was too short to form the kind of community in which people create many active discussions. We need to conduct a longer-term experiment to evaluate the usefulness of the system for discussion.

COMMUNITY VIEWER: Visualizing Community Activities

COMMUNITY VIEWER (Nishimura et al. 1998) provides the following three types of information: (1) personal profiles of conference participants and their interests; (2) static relations among the participants, as derived from the personal profiles; and (3) dynamic activities of the participants, including the ongoing process of community formation. The goal of COMMUNITY VIEWER is to support the formation of communities wherein people are willing to share their personal interests. At the first stage of community formation, personal profiles can help people in deciding whom to share personal interests with. Sharing knowledge (to know what other people know) and activities (to do what other people do) is also useful. Thus, the question is how to increase mutual interests knowledge, and activities without infringing on people's privacy. This dilemma became a serious problem in designing COMMUNITY VIEWER because private communication takes a more important role in community support systems than in traditional groupware.

These types of information are provided in a unified framework called the party room, which is a virtual place for visualizing the activities of the community. To enhance the awareness of the community among the participants and protect their privacy, we introduced the new concept of reflector icon. We can observe the

dynamic activities in the community by glancing over the room. Our party room has two display modes: (1) the overview mode that takes in the whole room and (2) the detailed mode that observes the activities of each individual. Figure 9a shows a typical screen image in the detailed mode.
[Figure 9 ILLUSTRATION OMITTED]

Each icon can represent abstract information about personal activities without infringing on personal privacy. In the party room metaphor, the reflector icon is realized as a face mark (figure 9). One icon represents one participant of the community. In COMMUNITY VIEWER, people can easily access personal profiles by selecting reflector icons. A detailed profile is shown in figure 9b.

Several functions are provided to distinguish among individuals in the party room. Suppose you want to find those who work in research fields similar to yours. You can do this by specifying the appropriate keywords as search conditions. When the system finds the other participants who match the conditions, the color of the corresponding reflector icons change, and the icons move toward your icon. As a result, you can easily access the personal information as needed and determine to whom to send a meeting request.

The ongoing interactions in the community are visualized in the party room. The behavior of each icon reflects the current activity of the corresponding person in the following manner: In the default state, when a user is not interacting, his/her icon walks toward a randomly selected table to sit down in the party room. After staying at the table for a time, the icon leaves for a different table. When one interacts (sends a message, reads a personal profile, and so on) with others in the virtual party room, the corresponding icon behaves as if approaching and talking to others. As the degree of interaction increases, icons start to congregate, mirroring what is seen at parties. The length of time over which the icons associate is proportional to the degree of interaction between the people. Thus, users get the impression that people in the community are actively interacting, although the contents of the interactions are not displayed.

To realize this function, the system has to judge who is interacting with whom and to what degree. In the experiments, we used the number of accesses to the personal profile as the degree of interaction. In general, e-mail, telephone calls, meeting arrangement requests, and any other interaction can be considered to indicate the degree of interaction. Each PDA calculates how much its owner is interested in others and sends the information to the community server. The server accumulates all the information and sends a summary back to each PDA. To visualize personal characteristics in a community, the reflector icon changes its size according to run-time statistics. People can know who is attracting attention and who is currently active by observing icon size.

Interaction Analysis

The challenge in this section is to investigate quantitatively the role of mobile computing in an actual international conference. To analyze a large amount of noisy data, we were careful in selecting key data items. Our strategy for analysis of log data was to select data items that were independent of the provided software functions and implementations. In addition, we excluded the log data of project members and system operators to get accurate use results from true participants.

Our first goal was to determine how mobile computing is used in conferences and how its use differs from that of desktop computing. Figure 10 represents the overall activities during the conference. The x-axis indicates the hour, and the y-axis indicates the number of events within this hour. The first two days (9-10 December) were for workshops and tutorials, and the next three days (11-13 December) were for technical programs. The reception was held on the evening of 11 December, and the excursion to Nara Park was set for the afternoon of 13 December.

[Figure 10 ILLUSTRATION OMITTED]

The following features can be observed in figure 10:

Continuous use: Except for the reception and excursion, the system was used continuously, even during technical presentations. People sometimes retrieved related information using PDAs while they listened to a presentation.

Midnight use: People used PDAs even after the conference, especially in their hotels after dinner. The highest peak in use, at midnight on 10 December, shows that PDAs are actively used just before the technical program starts.

Figure 11 shows the data for traffic on the third day, which showed the most typical traffic characteristic of the project. For purposes of comparison, it also shows the number of telephone calls made in one day on PASEO (Niwano et al. 1997), a multimedia communication market trial service provided by NTT Future Agent Network Inc., which also used MAGIC-LINK, the same PDA as in our experiment.

[Figure 11 ILLUSTRATION OMITTED]

In PASEO traffic peaked at 9:00 AM, noon, and 8:00 PM. These traffic peaks were because most users on PASEO had difficulty accessing the server during business hours. In the mobile assistant project, however, traffic dropped on the evening of the third day and the afternoon of the fourth day. These periods match the times for the conference reception and a postconference excursion, during which most users did not use their PDAs. Furthermore, the traffic peaked at midnight on each day Except during the reception and

excursion and in the very early morning, the system was used continuously. In this experiment, users could use their PDAs whether they had connected to the server, even during others' technical presentations. Traffic patterns on public services correlate with such things as hours of business, and the traffic pattern in this experiment correlates with the conference schedule. Thus, we learned that the use of a mobile computing system strongly reflects the users' situation.

Use of Services

We then analyzed the demands in international conferences by investigating how the services were used during the conference. Figure 12a describes the week's trend of participants' activity. The activity was low on 9 and 13 December because the PDAs were picked up and returned on these two days. At first glance, except for the first two days, activity was high at the beginning of the conference and gradually decreased toward the end. However, a structural trend exists as follows:
[Figure 12a ILLUSTRATION OMITTED]

Demand for e-mail services was fairly steady during the conference. The number of related events (sending and receiving messages) was not dependent on the progress of the conference but, rather, on the number of PDAs in use.

However, the demand for information services changed dramatically. Conference information services were used most on 11 December, the start of the technical programs. The demand for personal information services was high at the beginning of the conference and slowly decreased thereafter.

Although the number of accesses was not large, the demand for statistics feedback showed a trend similar to that of personal information services: active use at the beginning of the conference with a slow drop-off thereafter.

Figure 12b shows the daffy trend of participants' activity. The x-axis indicates three-hour blocks, and the y-axis indicates the sum of events that occurred within the corresponding time interval during the five days. The major results obtained from this figure are as follows:

[Figure 12b ILLUSTRATION OMITTED]

Activity increased from morning to night. Lunch time did not affect the users' behavior, but dinner time (reception and excursion) substantially decreased their activity. A remarkable peak appears from dinner to midnight. E-mail services were used steadily all day. Messages from the internet were received all night.

Conference information was retrieved frequently in the morning and after dinner, but personal information was often retrieved in the afternoon and particularly after dinner.

In summary, the demand for e-mail services was steady, but that for information services was heavily dependent on the conference structure. In previous conference support, desktop computing primarily provided e-mail services so that users could continue their business at the conference site. Mobile computing, however, can support more conference-oriented information services. The results of log data analysis create the picture of people using their PDAs after dinner in their hotel rooms to get additional information on other people whom they met in the afternoon and make an action schedule for the next day.

Connection Refusal

We also tried to find the effects of the transmission rate and communication errors on mobile computing services. For e-mail, figure 13 shows the data-transmission rate.

[Figure 13 ILLUSTRATION OMITTED]

In figure 13, points marked with an asterisk mean values for which the transmission had an incomplete status because of errors or user interruptions. In figure 13, the most important characteristic is the distribution of the data-transmission rate. Typically, the data-transmission rate on services using a wired local area network follows a normal distribution. The data-transmission rate in this experiment, however, has an offset distribution. Almost all measured points are concentrated around the line that defines an average data-transmission rate of 400 bits/s. Some points appear on the higher-rate side, but a few points appear on the lower-rate side. An average data-transmission rate of 400 bits/s is low compared with the 9600 bits/s of the physical layer on our mobile computing system. From these results (the low transmission rate and concentrated distribution), we estimate that in most cases of transmission in this experiment, the phenomena that caused a low-rate of transmission were transmission delay, bit error on the transmission layer, the retransmission of packets, and so on.
In questionnaires, users claimed that the data-transmission rate of communication with the server system was slow (transmission rate = 400 bits/s and setup time = 29 s; negotiation between moderns = 23 s, and the time to establish PPP [point-to-point protocol] = 6 s).

Figure 13 shows that user interruptions occurred during the early stages of transmission. In this experiment, the total size of data to be transmitted and the amount transmitted to this point were shown on the client system. Thus, users were able to estimate the total time of transmission. If the user knew that transmitting data from the server would take a long time, he/she often interrupted the

transmission.

Conclusions

Through the ICMAS'96 Mobile Assistant Project, we provided digital mobile assistants as agents for community support. The system was actually tested at an international conference, and the data collected are being analyzed.

The analysis of user behavior clearly illustrates the role of mobile computing in supporting international conferences. PDAs were continuously used not only at the conference site but also elsewhere after dinner. People used PDAs more often than desktop computing but for short periods only E-mail services were used steadily and independently of the conference structure, but the peak of information services depended on the progress of the conference.

Important for the design of information services on mobile computing we also found that users had a tendency to break the connection when they were forced to wait a long time for data transmission. Most users who had some error or interruption in transmission reconnected to the server system after a relatively long time.

We know that a single trial is not enough to obtain general conclusions. However, it is also true that large-scale experiments are not easy to repeat. To the best of our knowledge, no analytic reports have been published on conference support services, even for desktop computing. We believe it is valuable to share the results of our log data analysis with other researchers who might perform similar experiments in the future.

Postquestionnaire data showed that our trial was considered interesting, although people were not fully satisfied with the PDAs and the services provided. (See the sidebar for more detailed comments from participants.) The participants showed a deep interest in mobile computing for convention support. This project was planned as a step toward a new research initiative on communityware (Ishida 1998).

Systems that are to be used in a community are more difficult to design than point-to-point communication services. They cannot be evaluated by the simple sum of the impression of the users. The effectiveness of such systems must be verified through experiments involving large numbers of people. Through real experiments, we expected to confirm what the role of mobile computing in social interactions among human communities is. Figure 14a indicates the degree of satisfaction with the provided services, and Figure 14b indicates the degree of agreement with this trial. It is notable that the project was widely accepted, although we could not completely satisfy all participants. People claimed that the terminal was heavy and inconvenient to carry, the speed of communication was slow,

and so on. However, we can see that the users showed a deep interest in mobile computing for community support.
[Figure 14 ILLUSTRATION OMITTED]

Acknowledgments

The ICMAS '96 Mobile Assistant Project was conducted collaboratively by NTT Information and Communication Systems Laboratories, Kyoto University, Nara Institute of Science and Technology, and Kobe University and in cooperation with NTT Kansai Mobile Communications Network Inc., General Magic Inc., NTT Multimedia Service Promotion Headquarters, and Keihanna Interaction Plaza Inc. We want to thank Katsumi Tanaka, Shiro Takada, Rie Ohtsubo, and Katsumi Takahashi for their help in this project. We would like to thank the participants of this project for their cooperation and encouragement.

Notes

(1.) A more detailed project report can be found in Nishibe et al. (1998).

(2.) Now, agent-oriented mobile computing languages are mainly based on JAVA, for example, AGLETS and ODYSSEY.

References

Bly, S.; Harrison, S.; and Irwin, S. 1993. Media Spaces: Bringing People Together in a Video, Audio, and Computing Environment. Communications of the ACM 36(1): 28-47.

Engel, J. F.; Blackwell, R. D.; and Miniard, P. W. 1990. Consumer Behavior. 6th ed. Fort Worth, Tex.: Dryden.

Gaines, B. R., and Shaw, M. L. G. 1994. Using Knowledge Acquisition and Representation Tools to Support Scientific Communities. In Proceedings of the Twelfth National Conference on Artificial Intelligence, 707-714. Menlo Park, Calif.: American Association for Artificial Intelligence.

Ishida, T., ed. 1998. Communityware: Concept and Practice. New York: Wiley.

Kraut, E. R,; Scherlis, W.; Mukhopadhyay, T.; Manning, J.; and Kiesler, S. 1996. HOMENET: A Field Trial of Residential Internet Services. In Proceedings of the International Conferences on Human Factors in Computing Systems, 284-291. Vancouver, B.C., Canada: International Conferences on Human Factors in Computing Systems.

Maeda, H.; Kajihara, M.; Adachi, H.; Sawada, A.; Takeda, H.; and Nishida, T. 1997. Weak Information Structures for Community

Information Sharing. International Journal of Knowledge-Based Intelligent Engineering Systems 1(4): 225-234.

Mynatt, E. D.; Adler, A.; Ito, M.; and O'Day, V. L. 1997. Design for Network Communities. In Proceedings of the International Conferences on Human Factors in Computing Systems, 210-221. Atlanta, Ga.: International Conferences on Human Factors in Computing Systems.

Nishibe, Y.; Morihara, I.; Hattori, F.; Nishimura, T.; Ishida. T.; Maeda. H.; and Nishida, T. 1998. Mobile Digital Assistants for International Conferences. In Communityware: Concept and Practice, ed. T. Ishida, 245-284. New York: Wiley.

Nishida, T.; Takeda, H.; Iino, K.; and Nishiki, M. 1995. A Knowledge Media Approach to Ontology Development. In Toward Very Large Knowledge Bases: Knowledge Building, Knowledge Sharing, ed. N. J. 1. Mars, 84-94. Amsterdam, The Netherlands: IOS.

Nishimura, T.; Yamaki, H.; Komura, T.; Itoh, N.; Gotoh, T.; and Ishida, T. 1998. COMMUNITY VIEWER: Visualizing Community Formation on Personal Digital Assistants. Paper presented at the 1998 ACM
Symposium on Applied Computing (ACM SAC '98), Atlanta, Georgia. Forthcoming.

Niwano, E.; Okamoto, K.; Sakanoue, K.; Otaka, H.; Katsumata, M.; and Maeda, K. 1997. Evaluation of Large-Scale Mobile Multimedia Communication System Built on Telescript Agent Platform--The PASEO Service. In Proceedings of IEEE Multimedia 97. Washington, D.C.: IEEE Computer Society.

O'Day, V. L.; Bobrow, D. G.; and Shirley, M . 1996. The Social-Technical Design Circle. In Proceedings of the International Conference on Computer-Supported Cooperative Work, 160-169. Boston, Mass.: International Conference on Computer-Supported Cooperative Work.

Ohtsubo, R.; Takahashi, K.; Nishibe, Y.; and Morihara, I. 1998. ACTION NAVIGATOR: An Information Service Based on Agent Communication for Supporting Decision Making. In Proceedings of the Third International Conference on the Practical Application of Intelligent Agents and Multiagent Technology (PAMM-98). London: International Conference on the Practical Application of Intelligent Agents and Multiagent Technology.

Olson, J. S.; Olson, G. M.; and Meader, D. K. 1995. What Mix of Video and Audio Is Useful for Small Groups Doing Remote Real-Time Design Work? In Proceedings of the International Conferences on Human Factors in Computing Systems, 362-368. San Francisco,

Calif.: Morgan Kaufmann.

Sproull, L., and Kiesler, S. 1988. Reducing Social Context Clues: Electronic Mail in Organizational Communication. In Readings in Computer-Supported Cooperative Work, 684-712. San Francisco, Calif.: Morgan Kaufmann.

Weiser, M. 1993. Some Computer Science Issues in Ubiquitous Computing. Communications of the ACM 36(7): 74-83.

White, J. 1994. Telescript Technology: The Foundation for the Electronic Marketplace, White Paper 1994, General Magic, Sunnyvale, California.

RELATED ARTICLE: Comments from Participants

* It is a great initiative and I think it is the basis for the future mobile agent. I sometimes found it cumbersome; in general, if the communicational aspects improve, it will be of great value.

* It'll be more interesting if the participants use it more often. Less info, less news, less posting is not interesting.

* In general, the project is interesting and seems to have many possibilities. The concrete realization has many difficulties and problems. Besides those already mentioned, there have to be people doing first steps, i.e., posting news, inviting other people to discussion, meetings, etc. Without proving this initial start by yourselves, I do not think that much action will happen (I am also to blame, but it seems that nobody did anything, so I am not the only one.)

* I think the project is very interesting and potentially very important. The software should be enhanced, as well as the quality of the information, and the way it could be modified. The design of the user interface should be changed; 1-15 not very clear how to use it.

* Unfortunately the interaction speed/ease was far too slow for me to use. It was much easier just to meet people and I could never get e-mail to work. I did not have time to visit the help desk to see what was wrong. Good idea but the speed wasn't useful for me.

* The project was nicely arranged and I appreciated being a part of it. I was very disappointed with the PDA for several reasons. First, I had to have a special address that people at home did not know about. Second, I really, really, missed standard tools like TELNET and FTP. Finally, If you were taking notes on talks, the only thing you could do with them was send them home by e-mail. It would have been nice to be able to create files that could be transferred, through FTP, to your home server.

Yoshiyasu Nishibe is a research engineer at NTT Information and Communication Systems Laboratories. He received his B.E. and M.E. in 1987 and 1989, respectively, from the Department of Electrical Engineering, Shinshu University. His research interests include agent-communication services, multiagent system, communityware, and internet services. His e-mail address is nishibe@isl.ntt.co.jp.

Hiroaki Waki is a research engineer at NTT Information and Communication Systems Laboratories. He received his B.E. and M.E. in 1989 and 1991, respectively, from the Department of Information and Computer Sciences, Toyohashi University of Technology. His research interests are agent-communication services, mobile computing, and real-world-oriented interfaces. His e-mail address is waki@isl.ntt.co.jp.

Ichiro Morihara is a general manager of Kansai Systems Development Center at NTT Kansai Regional Headquarters. He received his B.E. and M.E. from Kyoto University in 1978 and 1980, respectively. His research interests include knowledge engineering, agent-communication systems, and mobile computing. His e-mail address is morihara@sdc.bch.ksi.ntt.co.jp.

Fumio Hattori is the executive manager of Computer Science Laboratory, NTT Communication Science Laboratories. He received his B.E. and M.E. in 1973 and 1975, respectively, from the Department of Electrical Engineering, Waseda University. His research interests include intelligent systems, software agents, multiagent systems, and communityware and social interaction. His e-mail address is hattori@eslab.kecl.ntt.co.jp.

Toru Ishida is currently a professor in the Department of Social Informatics at Kyoto University. He wrote Parallel, Distributed, and Multiagent Production Systems (Springer-Verlag), Real-Time Search for Learning Autonomous Agents (Kluwer), and Communityware: Concept and Practice (Wiley). His e-mail address is ishida@kuis.kyoto-u.ac.jp.

Toshikazu Nishimura is currently an assistant professor in the Department of Social Informatics at Kyoto University. He received his B.E. and M.E. in 1990 and 1992, respectively, from Kyoto University. His research interests include computer-mediated communication and mobile computing systems. His e-mail address is nisimura@kuis.kyoto-u.ac.jp.

Hirofumi Yamaki is currently a graduate student in a Ph.D. program in the Department of Information Science at Kyoto University. He received his B.E. and M.E. in 1995 and 1996, respectively, from the Department of Information Science at Kyoto University. He has been working on a project that aims to create systems that support a community of humans using a computer network. His e-mail address is yamaki@kuis.kyoto-u.ac.jp.

Takaaki Komura is currently a graduate student in the Department

of Information Science at Kyoto University. He received his B.E. from Kyoto University in 1997. He has engaged in research in information-processing multi-cast technology and layered encoding of multimedia data. His e-mail address is emon@kuis.kyoto-u.ac.jp.

Nobuyasu Itoh is currently a student in the Graduate School of Human and Environmental Studies at Kyoto University. He received his B.E. from Kyoto University in 1997. He has been studying medical informatics and working as a member of the Sensible Human Project. His research interests include applying computer science to health care. His e-mail address is nobuyasu@kuis.kyoto-u.ac.jp.

Tadahiro Gotoh is currently a research scientist in the Communications Research Laboratory Ministry of Posts and Telecommunications, Kashima, Japan. He was working on the ICMAS '96 project during his stay at the Department of Information Science at Kyoto University in 1996. He has engaged in research in time synchronization through computer networks. His e-mail address is tara@crl.go.jp.

Toyoaki Nishida is a professor in the Artificial Intelligence Laboratory at the Graduate School of Information Science at the Nara Institute of Science and Technology. He received his B.E., M.E., and doctor of engineering from Kyoto University in 1977, 1979, and 1984, respectively. His research interests cover knowledge media, agent technology, knowledge sharing, and qualitative reasoning. He is an area editor (intelligent systems) of New Generation Computing, an editor of Autonomous Agents and Multiagent Systems, a fellow of the Foundation of Physical Standardization Agents, and a trustee of the Japanese Association for Artificial Intelligence. His e-mail address is nishida@is.aist-nara.ac.jp.

Hideaki Takeda is an associate professor in the Artificial Intelligence Laboratory at the Graduate School of Information Science at the Nara Institute of Science and Technology. He received his Ph.D. in precision machinery engineering from the University of Tokyo in 1991. He has conducted research on a theory of intelligent computer-aided design systems, in particular experimental study and logical formalization of engineering design. He is also interested in multiagent architectures and ontologies for knowledge base systems. His e-mail address is takeda@is.aist-nara.ac.jp.

Atsushi Sawada is an associate professor in the Data Processing Center at Kyoto University. He received his B.E., M.E., and doctor of engineering from Kyoto University in 1990, 1992, and 1997, respectively. His research interests cover software-engineering environments and information modeling and sharing for software development. His e-mail address is sawada@kuis.kyoto-u.ac.jp.

Harumi Maeda is a JSPS Research Fellow and a Ph.D candidate in the Artificial Intelligence Laboratory at the Graduate School of

Information Science at the Nara Institute of Science and Technology. She received a B.A. in psychology from Kyoto University in 1986 and an M.Sc. in computation from the University of Manchester Institute of Science and Technology in 1994. She worked for Fujitsu Ltd. as a systems engineer from 1986 to 1995. Her research interests include knowledge representation, human factors in computation, and creative thinking support. Her e-mail address is harumi-m@is.aist-nara.ac.jp.

Masao Kajihara is an M.E. student in the Artificial Intelligence Laboratory at the Graduate School of Information Science at the Nara Institute of Science and Technology. He received his B.E. in 1996 from the Department of Information Science at Kyoto University. His research interests include augmentation of human memory, knowledge representation, and multiagent systems. His e-mail address is masao-k@is.aist-nara.ac.jp.

Hidekazu Adachi is an M.E. student in the Artificial Intelligence Laboratory at the Graduate School of Information Science at the Nara Institute of Science and Technology. He received his B.E. from the Department of Environmental and Sanitary Engineering, Kyoto University, in 1996. His research interests include ontologies, computer-aided design, and text analysis. His e-mail address is hidekaa@is.aist-nara.ac.jp.

SearcherOct 1, 1998

Searcher Responsibility for Quality in the Web World.

Author/s: T.R. Halvorson

T.R. Halvorson is the author of Legal Liability Problems in Cyberspace: Craters in the Information Highway (Houston, Texas. Burwell Enterprises, Inc., 1998), and proprietor of Pastel Programming, producers of specialized software for searchers.

What would be the word for it? Tepid? Ambivalent? Unimpressed? I would venture to say that the private attitudes of professional searchers toward the Internet are not as enthusiastic as those of the general population, or even as enthusiastic as our own public statements.

When commenting publicly, we concern ourselves with credibility. We wonder how much of Internet hype people believe. True, the consumer press now includes cautionary material. Still no info pro dares criticize the mighty Net too publicly, lest we cross conventional wisdom and sound out of date or like sour grapes. Perhaps we do envy committed Netters. What if other people really do experience the Web better than we do? We would not want anyone to think we are info-Victorians in the Brave New Information World.

Do we unconsciously worry that if we do not acquiesce in the hype, the perceived need for us might diminish? We have long yearned for wider markets and a deepened appreciation for search intermediaries in existing markets. And here comes the big breakthrough -- the Web. But why did the Web have to arouse everyone's expectations, and then fail to perform effectively while still diminishing, if not destroying, our pre-Web online performance tools?

Whatever the reasons, while everyone else is clicking on the Web, something is not clicking with many information professionals. We keep coming home to traditional database services. We experience self-doubt about our Web-avoidant behavior. Relief has started to come, but not in our Internet searches. Instead some professional

seminars have turned into therapy sessions: "Hello. My name is T. R. and I am a Web-avoidant searcher."

If you think you can get away with Web-avoidance, think again. The Law may not let you. The Law? That's right, the law. Your clients may now have legal duties to search the Web. As their agent, this means you can avoid Web searching no longer.

The Law Weighs In

In legal circles, Whirlpool Financial Corp. v. GN Holdings, [Inc.sup.1]. has been widely discussed as imposing a legal duty to search the Internet. In July 1991, Whirlpool Financial Corporation made a $10 million loan to GN Holdings, Inc. to partially finance GN's purchase of Cross Country Healthcare Personnel, Inc. CCHP provided hospitals across the United States with temporary health care personnel such as nurses, physical therapists, and occupational therapists. To solicit financing for the purchase, GN Holdings, with the help of Shearson Lehman Brothers, circulated a Private Placement Memorandum, which provided narrative information, historical financial data, and projections for future performance. Later, before Whirlpool purchased the note, these projections were "revised and made significantly less optimistic." The Private Placement Memorandum and the subsequent amendment set out projections for, among other things, net sales, operating profit, pretax profit, and net income during the remainder of 19 91 and fiscal years 1992 and 1993. To allow Whirlpool to monitor GN's performance, the loan agreement required GN to send Whirlpool monthly, quarterly, and yearly financial statements accompanied by management reports. GN complied with these requirements and Whirlpool received the financial statements as scheduled.
The projections painted a much rosier picture than what actually unfolded. In the years 1991 through 1993, net sales fell 32 to 48 percent lower than projected, operating profit 50 to 73 percent lower, pretax profit 95 to 257 percent lower, and net in come 104 to 283 percent lower. On September 12, 1991, shortly after receiving the first financial statement, Whirlpool ac count executive Steven Furman traveled to Boca Raton, Florida (GN's headquarters) to question GN executives about the discrepancies between the financial projections and actual performance. Furman made similar trips on February 5, 1992 July 24, 1992; October 16, 1992; and June 23, 1993. At these meetings, GN executives explained that the general economic recession had softening the demand for traveling nurse services, but that better times were just around the corner.

Happy days did not arrive, and GN defaulted on the interest payment due Whirlpool on April 1, 1994. Whirlpool Financial Corporation filed a securities lawsuit on July 11, 1994, against GN Holdings, Inc., and W. R. Grace & Co.. seeking rescission of Whirlpool's $1C million loan to GN. In its complaint, Whirlpool alleged, among other things, a violation of Section 10(b) of the Securities Exchange Act of 1934 and a violation of Section 12(2) of

the Securities Act of 1933. The federal district court dismissed Whirlpool's federal claims with prejudice because they were filed beyond their respective statutes of limitation. Whirlpool appealed to the U. S. Court of Appeals for the Seventh Circuit.

On appeal, Whirlpool's brief summarized its view of the significant factual underpinnings of its claim:

"The Revised Projections were unreasonable when made because they were based on GN's historical performance, while defendants knew, but did not disclose, that the industry was subject to material adverse trends, including proposed and enacted state and federal legislative and regulatory actions which would limit GN's revenue, increased competition to recruit nurses, trends toward reduced usage of temporary nurses, trends toward increased compensation and benefits for traveling nurses, shifts from in-patient to ambulatory care, and trends toward hospital closings and a decrease in licensed beds."

The Court of Appeals said the basis of Whirlpool's claim was that the defendants' revised projections were made without a reasonable basis in light of the adverse legislative, regulatory, and industry trends known to them. Such a Rule 10b-5 claim based on projections made in bad faith or without a reasonable basis is cognizable under the securities laws.

In both Section 12(2) and Rule 10b-5 actions, a plaintiff must file its claim for relief within one year from the time that its action accrues. Further, inquiry notice suffices to start the limitations clock ticking. Inquiry notice starts the running of the statute of limitations when the victim of the alleged fraud became aware of facts that would have led a reasonable person to investigate whether he might have a claim:
"[T]he information Whirlpool says it needed to 'uncover' the alleged fraud was in the public domain. In today's society, with the advent of the 'information superhighway,' federal and state legislation and regulations, as well as information regarding industry trends, are easily accessed. A reasonable investor is presumed to have information available in the public domain, and therefore Whirlpool is imputed with constructive knowledge of this information. See Eckstein v. Balcor Film Investors, 58 F.3d 1162, 1169(7th Cir. 1995).

"When examined consistent with an objective reasonable diligence standard, the only reasonable inference that could be drawn from the facts as alleged by Whirlpool was that it was put on inquiry notice before July 11, 1993. Thus, the district court properly determined that Whirlpool's federal claims, not asserted until its complaint was filed on July 11, 1994, were time barred."

Whirlpool tried to meet inquiry with equitable estoppel. It said, the mouths of the defendants should be stopped from arguing the

statute of limitations because they "lulled" Whirlpool into missing the statute of limitations. Whirlpool contends that the explanations Furman received for GN's poor performance - primarily the general economic recession concealed he defendants' alleged fraud. Applying he law of the statute of limitations would, thus, be inequitable. Equitable estoppel may apply where a defendant took active steps to conceal evidence from the plaintiff that the plaintiff needed in order to determine it had a clam. In the court's view, however, this also is answered by the Internet:

"However, in light of our determination that the information Whirlpool needed to uncover the fraud was in the public domain, GN's continued attempts to explain away the discrepancies between the revised projections and the actual earnings could not have prevented Whirlpool from filing its complaint on time."

The court is unsympathetic and says, in effect, the information is on the Internet, "deal with it."

The legal system appears poised to impose the same standard upon itself. The Pennsylvania Bar Association publishes a series of malpractice avoidance articles at its Web site. One of them, entitled "Surfing the Internet - It's Sink or Swim," argues that attorneys must use the Internet to search the law and investigate facts. They cite an unreported trial court ruling in Wallach v. Stradley Ronon Steven & [Young.sup.2]. The plaintiff, coming off his victory in a previous lawsuit, sued his adversary's lawyers. He claimed they were liable for malicious prosecution because they had an inadequate factual basis for their decision to file the earlier lawsuit. Traditionally attorneys may rely upon their clients to supply the information forming the basis of the decision to file a lawsuit. The lawyers raised that defense, saying that at the stage of filing the suit, they relied on their client. The article purports that the court's ruling modified the traditional rule. The court's ruling made it plain that relying on the client may not suffice. At least where lawyers have some indication that a client's information might be incorrect, the lawyers must perform an independent investigation to clear their liability for malicious prosecution.
Other professional literature contends that failure to use computerized legal research constitutes legal [malpractice.sup.3].

"We have already seen online data bases, such as LEXIS and WESTLAW, become standard fixtures in our offices. As the burden of performing factual investigation is increasingly placed upon attorneys, effective use of the computerized legal research has become a requirement of our due care.... In the same way, use of the Internet will become a required aspect of every attorney's arsenal."

The article then nicely positions professional searchers as indispensable members of the legal team:

"Realizing that not every attorney can master the Internet, and, perhaps, that no one attorney can master the entirety of the Internet, we will need to integrate information professionals into our day-to-day practices, even more so than litigation firms have been utilizing private investigators."

If these cases and articles truly represent trends and not anomalies, neither clients nor lawyers can avoid searching the Internet. Perhaps professional searchers cannot avoid the Web either.

Quality Fears

When we track the court's language and reasoning, we can see that it may not even suffice to run these searches on DIALOG or LEXIS-NEXIS. Federal and state legislation and regulations have always resided in the public domain in the copyright sense, but not in the sense of easy access. In a single sentence, the court ties ease of access to "today's society" and "the advent of the 'information superhighway.'" And the court also places the duty on the "reasonable investor," not just lawyers with special professional duties to search LEXIS-NEXIS or Westlaw, not just large corporations with librarians on staff and Dialog accounts. The court does not charge the reasonable investor with a duty to search Dialog or LEXIS-NEXIS. Instead, it appears to expect the reasonable investor to find whatever that $19.95, flat rate, unlimited Internet access could find (in the language linking "easy access" and the "advent" of the "information superhighway"). The court refutes Whirlpool's claim of the great burden imposed to "u ncover" the information, by saying the information is "easily accessed" on the Internet. The expensive sources you may search if you will; the "free" ones you ignore at your peril.

It doesn't seem fair. Why not? Because Internet information and Internet searching tools often lack sufficient quality to find the right stuff reasonably. Professionals know the unreliability of the conventional wisdom that "everything is on the Internet. Just tap a few keys on your computer and out comes everything."

Why do many information professionals become Web-avoidant? Because they do not have enough authority over Net information to match the responsibility imposed to find it. When people talk about command language or command mode on traditional services, they mean more than computer jargon. The word "command" carries a latent psychological meaning. It suggests that the searcher has authority over the database. On the Web, most searchers do not know what Web search engines really do and do not trust their ability to make the engines do what they want. Betsy Anagnostelis, Alison Cooke, and Alison McNab have described some of the problems of searching the Web in an article entitled "Thinking Critically about Information on the [Web:.sup.4]"
"A search engine may return thousands of sites with very little descriptive information to guide the user, who must consequently

expend time and energy to follow the links suggested, and sift through the results. Following the links alone can lead the user down numerous blind alleys and dead ends as sites move, or sites turn out to offer material of interest or relevance or to contain large graphics which take so long to download that the user loses interest. Medical information on the Internet today, for example, often has questionable accuracy with top quality sites sitting side by side with dubious [material.sup.5]."

No data, the wrong data, bad data, and unreliable data cause problems for client and searcher alike. It might even open the searcher to legal liability.

All sorts of definitions of quality have come forward. Here's another, which I consider particularly appropriate to the Web's quality quagmire: Quality is the combination of traits of information which yields to searchers sufficient authority over the information to match a professional searcher's responsibility for finding it.

The SCOUG Rating Scale

The earliest full-orbed view of quality and value of information in the electronic age I could find came out of the 1990 annual retreat of the innovative Southern California Online Users Group [(SCOUG).sup.6]. From that effort came the SCOUG Rating Scale. The Scale established a framework for judging performance in 10 broad categories:

* Consistency

* Coverage and Scope

* Timeliness

* Accuracy/Error Rate

* Accessibility/Ease of Use

* Integration

* Output

* Documentation

* Customer Support and Training

* Value-to-Cost Ratio

Look at the Web through SCOUG Rating Scale lenses and you find a medium with serious quality problems, even for the best of

databases. Let's try to apply the SCOUG Rating Scale to Web versions of a familiar database, like MEDLINE. Searchers can find free Web access to MEDLINE on several venues: Healthgate, Healthworks, Infotrieve, Medscape, Avicenna, NLM's Internet Grateful Med, NLM's PubMed, BioMedNet, and UKLON. You can get paid access to MEDLINE over the Web through many other venues. Betsy Anagnostelis and Alison [Cooke.sup.7] report on several ongoing comparative studies of the quality of MEDLINE services on the Web. Most of these studies combine criteria from various sources, and most use the SCOUG Rating Scale as their root. The studies have interest not only in their analysis of how the various Web versions of MEDLINE stack up, but also for how SCOUG-inspired rating systems fare when applied to Web services.

As part of the Organising Medical Networked Information (OMNI) [project.sup.8], information professionals devised a set of evaluation criteria, based primarily on the SCOUG [standards.sup.9]. Anagnostelis and Cooke presented the paper at Online Information 97, the 21st International Online Information Meeting, in London, December 1997. The paper, available online (see note 9), provides detailed notes on the use and implementation of the criteria. In evaluating the utility of the SCOUG Scales, the authors conclude: "[T]he proposed criteria enable detailed consideration of the range of features and facilities available from different services and were therefore found to be useful in evaluating a range of WWW-based MEDLINE services. The name 'MEDLINE' automatically offers credence to a service. However, using the criteria it has been possible to highlight the differences between services providing access to the 'same' database, including the lack of features traditionally associated with effective searching via Medline, and a lack of adequate and context sensitive help information to assist the user.

"... Services supporting natural language query input present novel challenges, especially in the absence of full explanations of precisely how a query is analyzed and how the output is ranked. The impenetrability of such services raises important support issues as users undoubtedly find them attractive for quick searches but may be missing references which are central to their work.

Using the proposed criteria as a guide when examining different versions of the same database, librarians and other information professionals should be able to make an informed assessment of the relative value and usefulness of a range of services. The criteria are designed specifically for comparing services available via the World Wide Web."

Since the SCOUG rating scales seem to adapt nicely to evaluation of databases like MEDLINE on the Web, I have a proposal. Let us use the Web to conduct and publicize SCOUG-inspired quality evaluations of selected Web resources. We are the information professionals. We and our organizations should get to work. We

should roll up our sleeves and start the testing. When Web products need improvement, we should have the evidence in hand and shout our message to vendors and content providers.

Studying the Wider Web Quality Problem

Applying the SCOUG Rating Scale to familiar databases like MEDLINE is relatively straightforward. Already familiar with the files, we only have to adjust the Scale to the many Web versions. Applying the Scale to the Web's own unique tools is more of a problem. How does one apply the SCOUG Rating Scale or any other set of quality criteria to Internet casinos, peep shows, family albums, bulletin boards, role playing games, and time interval digital camera shots out the office window of a university professor in Norway? Not immediately obvious, is it?

One approach is to view the Web as several spheres jumbled together, somewhat overlapping yet distinguishable. David Siegel has done that in "The Balkanization of the Web." He depicts the Web in a Venn diagram, a cross-section not an overview, with three circles: information, exchange, and entertainment.'0 Many sites have elements of all three; still, one can distinguish each element. Sometimes clients may request "information" found in the entertainment or exchange spheres. Serious retrieval problems may occur, but the unique content, unavailable on traditional database systems as a rule, makes dealing with those problems necessary. Quality problems come more into focus in what Siegel calls the information sphere. Here we can take a swipe at applying rating scales.

Some of the quality issues would raise eyebrows in any sphere. For example, in September 1997, HotBot added more computers to its array. They did this without taking the service offline. HotBot kept the database online, but not the complete database. Someone searching during that period of time would not know they only got hits from a subset of the database. That's not good service in anybody's venue.

Scope, Coverage, and Recall in Web Search Engines

Most of the popular search engines have acknowledged from the beginning that they never index CGI scripts that run queries, because they cannot. Those scripts assemble HTML pages dynamically in response to queries. The resources exist with the data in some form of database or collection on the Internet. But no page or presentation of the data exists until the script runs the against the database or collection. When run, the script creates an HTML page on-the-fly and HTTP transports it to my browser.

Another coverage problem stems from registration requirements. Sites like The New York Times bar users who have not filled out a registration form, while the Wall Street Journal Interactive Edition on the Web requires both registration and subscription payment.

Spiders, the software which indexes the Web for search engines, cannot easily register to view these sites. Usually registration pages and login boxes block the search engine spiders that crawl the Web from returning indexing information into their databases.

By the way, when it comes to monitoring all aspects of search engines, you can't go wrong with Danny Sullivan's Search Engine Watch," where I came across "Melee's Indexing Coverage [Analysis.sup.12]," which says:

"We began this project in response to the discovery that the AltaVista index retains only a sample of all the pages on medium to large sites. As the referenced page mentions, the folks at AltaVista are not denying this fact. On the other hand, if you look at AltaVista's home page, you sure don't see any hint that (a) their index is MUCH LESS than complete or that (b) their index has less coverage than other indices."

Bad enough, but look at what Melee offers to do:

"Wouldn't it be nice to have index machines that never missed any of the new pages on the Web? Wouldn't it be nice to have index machines that never had stale links? Well, the real world never will be that good. Melee's Indexing Coverage Analysis (MICA) examines the relative page coverage for a select group of search engines. Each week, Melee Productions will retest the engines on the list and publish an update to the MICA Report. We will be happy to test any publicly accessible search engine that supports date-range and host/domain constraints and purports to index at least one-fifth of the 'Web.' If your search engine is not on the list and you feel it should be, send us e-mail!"

They have weekly reports for prior weeks, but the week's report I looked at considered the HotBot search engine and the AltaVista search engine primarily. No other public search engines "that we know of both (a) indexes at least 20% of the Web and (b) supports combined date range and host/domain constraints on a search." Some services claim to have more than 90 percent of the Internet in their database, a total of almost eight million pages. However, if you read the fine print, you realize that they only indexed the text of about 1 million pages. For the other 7 million pages, their database only contains the link that point to the page.

Ninety percent, 20 percent, 50 percent - all arrived at by different types of counts and percentages. Greg R. Notess, Reference Librarian, Montana State University, performs periodic "apple-to-apple" comparative searches on nine leading Web search engines and presents the results in text, numbers, and graph form at his Web [site.sup.13]. He finds a marked disparity of recall. Perhaps more importantly, Notess produces a study of the degree of overlap in the hit sets - very little.

Ranking, Placement, and Precision

A multitude of Web-specific issues affect inclusion, exclusion, ranking, placement, and precision of Web search engine results. Searchers rarely understand the mysteries or esoteric semantic programming rules that prepare the offerings they see. Forces have arisen that work against the needs of searchers as the dominant influence in that programming.

Spam or its equivalent doesn't only occur in e-mail and newsgroups. Spain hits the Web search engines with static spam, opportunistic spam, trademark spam, blizzard text, metatag spam, ranking spam, and bridge pages. Netters fight back with legislation, spam penalties imposed by search engines, and by spam "narcing." Some approaches that search engines use to penalize spam do not discriminate sufficiently and can "protect" searchers from good information. Besides spam, business alliances, editorial partnerships, and pay-for-placement can affect search engine quality. Which link descriptions appear at the top of a list or oftenest may depend less on their relevance and occurrence patterns, than on a contractual relationship between the site or its representatives and the search engine company's management.

Currency

Currency is another Web quagmire. Browser caching on the client end and proxy caching and replication on the server end lead to the delivery of stale pages onto your screen. Technical studies give different estimates, anywhere from 50-90 percent, of the percentage of cache hits as opposed to transport of the page from the site of origin. MatchLogic, a leading online ad management firm, has produced a True-Count product which shows that the current methods of counting ad impressions typically underreport ad inventory usage by an average of 76 percent, and for the most popular pages the underreporting reaches 674 percent, thanks to [caching.sup.14]. The high percentage of cache hits would not matter if server caching software could maintain cache coherence, the synchronization of cached versions of Web pages with up-to-date versions at the sites of origin. Web coherence, however, is another quality topic for deep study.

Content Quality

Internet resources, in particular World Wide Web resources, continue to proliferate at an astonishing rate. Unlike professional journals and commercial publishers, who employ a system of editorial review and external referees to ensure the caliber of materials distributed, information can be spread over the Internet by anyone without regard to accuracy, validity, or bias. Due to its global structure, which encompasses a variety of legal systems and cultures, it is unlikely any one individual or nation can significantly influence, regulate, or change the chaotic state of flux that characterizes the World Wide Web. On the other hand, in my opinion, there is nothing

wrong with letting the Web serve as a venue for freedom, with letting it include the unscholarly and the trivial. The Web has purposes and functions that extend beyond our frustrations as searchers.

Nevertheless, a need exists for criteria and procedures that will assist students, teachers, scholars, and other users in evaluating the quality of Internet information and for standards to guide the design of Web resources. The University of Georgia has established a site to develop such a set of criteria and standards. Gene L. Wilkinson, Lisa T. Bennett, and Kevin M. Oliver of the Department of Instructional Technology at the University of Georgia have published a "Consolidated Listing of Criteria and Quality [Indicators.sup.15]." A panel of experienced Internet [users.sup.16] ranked the criteria.

Memorial University has published an interesting and very comprehensible evaluation scheme. Memorial's scheme neatly divides criteria into two main categories - concrete validity and context validity-with more detailed criteria underneath. As Memorial defines it, "concrete validity" focuses on validating items which should appear visible on a Web page itself (e.g., Webmaster contact information, author background descriptions, source information, etc.), while "context validity" can credit a page as reliable even when concrete validity is weak, e.g., when the page supplies links to supporting information and demonstrates no [bias.sup.17].

In an overly severe indictment of the Web's content quality problems, Matthew Ciolek does itemize real content quality [concerns.sup.18].

"The problems with the Web are many. WWW documents continue to be largely un-attributed, undated, and un-annotated. As a rule, information about the author and publisher is either unavailable or incomplete. Frequently, the rationale for placing a document online and information about how it relates to other materials is not explicitly stated, It has also been observed that the Web remains a place in which far too many resource catalogues seem to chase far too few original or nontrivial documents and data sets.

"Simultaneously, there are no commonly accepted standards for the presentation of online information. Instead, there is an ever-growing proliferation of publication styles, page sizes, layouts and document structures. Moreover, links to other Web resources tend to be established promiscuously, that is without much thought for the target's relevance or quality. There is also a pronounced circularity of links. This means that many Web pages carry very little information, apart from scantily annotated pointers to some other equally vacuous index pages that serve no other function apart from pointing to yet another set of inconclusive indices and catalogues. Finally, emphasis continues to be placed on listing as many hypertext links as possible - as if the reputation and usefulness of a given online resource depends solely on the number of Web

resources it quotes. In practice this means that very few such links can be checked and validated on a regular basis. This leads, in turn, to the frequent occurrence of b roken (stale) links.
"In such a vast and disorganized context, work on simple and low-content tasks, such as hypertext catalogues of online resources, is regularly initiated and continued at several places at once.... At the same time, more complex and more worthwhile endeavors, such as development of specialist document archives or databases, are frequently abandoned because of the lack of adequate manpower and funding."

There is an information swamp on the Web. Just look at the Kurt Vonnegut MIT commencement speech hoax, the Denny Reikert murder conspiracy hoax, the Chief Seattle eco-spiritualism speech hoax, virus hoaxes, financial hoaxes, journalism hoaxes, just to mention examples of one form of bogus information. Ciolek's work, with Irena M. Goltz, on the Information Quality WWW Virtual Library [http://www.ciolek.com/WWWVL-INfoQuality.html (URL last checked: August 28, 1998)] is priceless.

On the other hand, peer review, editorial policy, collegial confirmation, and other standard methods of assuring quality sometimes "protect" us from good information. One person's pigheaded stupidity might be another's brilliant insight. In my own profession, the law, Bernard J. Hibbitts, associate dean for Communications & Information Technology and professor of Law at the University of Pittsburgh School of Law, offers an example. He has skipped editorial review and self-publishes directly on the Web. In perhaps his best known article, "Last Writes? Re-assessing the Law Review in the Age of [Cyberspace.sup.19]," Hibbitts asserts that self-publishing offers academics the advantages of speed, freedom from student-editors' objectionable line editing, potential for publication of less traditional scholarly work, the opportunity to fully use the Web's hypermedia potential, and finally the benefit of later revision of scholarly work. He posted "Last Writes?" as Version 1.0. "Last Writes?" has received numerous si gnificant commendations and awards. Traditional law reviews sought to (re?)print it. One of the pages at Hibbitts' site lists where "Last Writes?" has been published, abridged, condensed, digested, reviewed or [linked.sup.20].

Information professional organizations like SCOUG, the Association of Independent Information Professionals (AIIP), SLA, etc., sought to synthesize this body of literature, adding their own contributions and drawing classic works in the more general literature on quality of information, and publish them with an eye to building sturdy, classic evaluation instruments. It is astonishing how sturdy the SCOUG Rating Scale has proved, especially considering that it was developed not iteratively, but during a one-time happening after which its creators rode off into the sunset, their perfect work complete. However now we need to apply its criteria and produce the measures. We ought to be more afraid of doing nothing than of

doing it wrong.

Discovering Web quality solutions

As the problems of the Web have proven maddeningly diverse, so are the solutions. Ciolek organizes the approaches into six categories:

1. programmatic approaches

2. procedural approaches

3. structuring approaches

4. bibliographical approaches

5. evaluative approaches

6. organizational approaches

Programmatic approaches to fixing the Web focus on providing online publishers a wide range of flexible tools for generation and manipulation of hypertext documents. Engineering approaches seek to construct interlocking modular, intelligent, or quasi-intelligent software agents that could organize, channel, and guide publishing and retrieval on the Web.

Efforts include the creation and enhancement of the Hypertext Markup Language (HTML), client-server software, creation of the PERL language, creation of the Java language, the Common Gateway Interface (CGI), and the creation of the Cascading Style Sheet (CSS 1) mechanism, XML, Style Sheet Languages, etc. One important work is the creation of the Platform for Internet Content Selection (PICS). This is a software tool to provide content labeling, rating systems, and access control to Web resources. PICS provides content providers a self-rating capability to describe and label the material they provide. PICS also provides for multiple third-party ratings by which independent labeling services could associate additional labels with content created and distributed by online authors.

The Dublin Core Metadata Element Set2 represents a reasonably potent yet simple method of enhancing the quality and retrievability of Web pages. Labels are defined for Title, Creator, Subject, Description, Publisher, Contributors, Date, Type, Format, Identifier, Source, Language, Relation, Coverage, and Rights. Searchers should promote such projects.

Important work on CGI and Java scripts could provide users with the means to create data-input pages for online collection of corrections, feedback, and other reader-supplied information. Searchers should look for applications that can increase quality in Web products, use

them and praise them loudly to encourage their proliferation.

Procedural approaches focus on raising the level of the practice of Web publishers. Often a publisher's proficiency in HTML, CGI scripts, and Web technologies constrains the quality of their Web sites and Web content. Procedural improvements can include HTML style guides, templates, presentation standards, redirection pages when URLs change, bandwidth conservation, and algorithmic sequences for data acquisition, data preparation, data formatting, document naming, directory naming, document installation, setting ownership and protection levels, updating indices and catalogs, and document maintenance.

Bibliographic approaches attempt to deal with the vagaries, transience, and morphology of Internet resources. Numerous citation formats have been proposed. Some scholars think that unless Web resources become commonly cited in academic discourse, they will be ignored and never improved. Of course, bibliographic control also motivates and accommodates archiving. Searchers should learn some schemes of bibliographic reference of Web materials and use them.

Evaluative approaches stem from the perception that Web resources, though dissimilar, share enough common characteristics with other resource to allow rating, grading, and scoring. The rating criteria developed at the University of Georgia discussed earlier in this article is just one example.

Organizational approaches hope that the information swamp can be drained and the ground placed into tamed cultivation by the energetic and competitive cooperation of individuals and institutions with a stake in the Web. The Internet ethic of sharing and collaborating inns through these approaches. The WWW Virtual Library Project serves as a typical example; others include the Clearinghouse for Subject Oriented Guides to the Internet and the Special Interest Networks (SINs). The SINs approach would combine the roles of information suppliers, distributors, and users.

Searchers of the world, unite! Let us get to work, like the information professionals we claim to be, and help make the Information Age work for us and our clients. It's our future. Let's control our fate.

Footnotes

1. 67 F.3d 605 (7th Cir. 1995).

2. Unreported cases have no standard form of legal citation. The article cites 31 Phila. (1996).

3. O'Connell, "Legal Malpractice: Does the Lawyer Have a Duty to Use Computerized Legal Research," 1991 (Nov/Dee) L.Prac.Mgmt.

23; Duane A. Daiker, "Computer-related Legal Malpractice: An Overview of the Practitioner's Potential Liability," 69 Fla.B.J. 12 (April. 1995). See also Robert C. Berring, "Technology and the Standard of Care for Legal Research,"

(3,4) Legal Malpractice Rep. 21(1992); Ola Najar, "Computerized Research in the Legal Arena: Developing a Standard of Research Sufficiency in Legal Malpractice and Rule 11 Actions," 39 wayne L. Rev. 1683 (1993); Mark J. Newman, "Attorney Research Malpractice," Practising Law institute Commercial Law and Practice Course Handbook Series PLI Order No. A4-4351, 590 PLI/Comm 11 (September 25, 1991); and Simon Chester, "Electronic Malpractice: Does Reasonable Competence Require Computer Research," 17(8) Law Prac. Mgmt. 23 (Nov/Dec 1991).

4. Anagnostelis, Betsy, Alison Cooke, and Alison McNab, "Thinking Critically about information on the Web" [http://omni.ac.uk/agec/vine.htm] (URL last checked: August 28, 1998)], document last modified Tuesday, 20-May-97 09:02:50 GMT.

5. Loc. cit., citing: (3) Uretsky, S., "Beware of Useless or Dangerous Medical Advice Online," Internet World, February 1996 [http://www.internetworld.com/ 996/02/medicine.htm] (URL last checked: August 28, 1998)].

6. Basch, Revs, "Measuring the Quality of the Data: Report on the Fourth Annual SCOUG retreat," Database Searcher, vol. 6, no. 8, October 1990, pp. 18-24.

7. http://omni.ac.ukJagec/iolim97/ (URL last checked: August 28, 1998).

8. OMNI Consortium, Medline on the Internet; last modified 5th May 1997, http://www.omni.ac.uk/general-info/internet medline. html (URL last checked: August 28, 1998).

9. Anagnostelis, Betsy, "Evaluation of Internet Medline for OMTI" http://www.omni.ac.ukigeneral-info/intmedeval

10. html (URL last checked: August28, 1998).

11. (URL last checked: August 28, 1998).

12. http://searchenginewatch.com (URL last checked: August 28, 1998).

13. http://www.melee.conijmica/uindex.htm] (URL last checked: August 28, 1998). http://www.onlineinccon1Jdatabase/OctDB97, netl0.html (URL last checked: August 28, 1998).

14. "ABC/MatchLogic Looking Beyond the Cache Barrier," Media Central, vol. 5, no. 228 (November 20, 1997). See also, "Resolving

Ad Undercounts May Boost Web Spending, Payoff," interactive PR & Marketing News, October 31, 1997; Rick E. Bruner, "MatchLogic Service Solves Cache Problem," Ad Age (October 13, 1997); and Evan Neufeld, "By Solving Proxy Caching, MatchLogic Opens a Pandora's Box of Impressions," Jupiter, October 13, 1997.

15. http://itechl.coe.uga.edu/Faculty/gwilkinson/rankings.html. (URL last checked: August 28, 1998).

16. The identity and qualifications of the members of the panel, and the method of selecting them, is set forth at http://itechl.coe.uga.edu/Faculty/gwilkinson/panel.html (URL last checked: August28, 1998).

17. http://www.stemnet.nf.ca/Curriculum/Validate/concrete.html (URL last checked: August 28, 1998).

18. Ciolek, T. Matthew, "The Six Quests for the Electronic Grail: Current Approaches to Information Quality in WWW Resources," Review Informatique et Statistique dons lea Sciences Humaines (RISSH), 1996, No. 1-4. Centre Informatique de Philosophic et Lettres, Universite de Liege, Belgium. pp. 45-71.

19. http://www.law.pitt.edu/hibbitts/lastrev.htm (URL last checked: August 28, 1998).

20. http://www.law.pitt.edu/hibbitts/owlw.htm and http://www.law.pitt.edu/hibbitts/oplw.htm (URLs last checked: August 28, 1998).

21. http://www.ub2.lu.se/tk/metadata/DC10cats.html (URL last checked: August 28, 1998).

Internet World Sept, 1998

SETTING UP SHOP.(electronic commerce storefronts for selling books were created using Yahoo Store and iCat Commerce Online) (Company Business and Marketing)

Author/s: Andrew Marlatt

A REPORT FROM THE TRENCHES

IW hits the Web, armed with widgets and a credit card. The mission: Build a business.

IF IT'S TOO GOOD to be true, the saying goes, it probably is, and the promise of building and running an Internet store "in minutes" using only a browser seemed to be just that.

After all, corporations commonly spend months and hundreds of thousands of dollars setting up sites. Could it really be possible for an Internet neophyte to create and run an attractive online store--and generate sales-for as little as $50 or $100 a month?

Internet World decided to find out. In early July, we set up two Web sites selling baseball-related books--one on Yahoo Store (formerly viaweb), and one on iCat Commerce Online (iCO), the two most prominent e-commerce services.

Reviewers have praised the simplicity of these services, which primarily target small and midsize businesses looking to get their Web feet wet.

But our goal was not to page through what these services have to offer. We wanted to build a full store, take orders, and track store activity, as well as seek out traffic by registering with search engines, setting up reciprocal links, and taking out banner ads. In short, we wanted do everything a real, relatively new-to-the-Web merchant would want to do.

First, we needed products to sell. We turned to The Scholar's

Bookshelf, a Cranbury, N.J.-based bookseller, and built our sites using titles from the company's Books on Baseball catalog. To ensure the sites received no special treatment, we never identified ourselves to iCat or Yahoo as being part of Internet World, and our calls to technical-support and sales staff were made only as storeowners. [We should also note that Yahoo Store provides statistics for Internet World's weekly "Commerce Quotient" feature.]

Before attempting to create the baseball book sites, we decided on some criteria Each store would sell the same 50 books--a number we settled on because both Yahoo Store and iCO charge on a sliding scale based on the number of items for sale, and 50 was a common cutoff point. To help build share in the online commerce market, ICO currently offers free space to clients selling fewer than 11 items, and charges $50 a month for up to 50 items. Yahoo Store has no free option, and it charges $100 a month for up to 50 items.

We planned to split the 50 books fairly evenly into five categories, such as "History" and "Teams." We also wanted left-side navigation, a search field on the home page, a white background with icons linking to each section, and the ability to use the "Books on Baseball" logo and scanned-in book covers. We would avoid plug-ins--after all, these turnkey services are designed specifically for novices--and do as little coding as possible.

Our other goals included making available a privacy statement and opt-out check boxes on order forms; a company info page with an executive message; and at least one page with our own content. For the latter, The Scholar's Bookshelf devised a trivia quiz whose answers referenced and hyper-linked to particular books. Finally, we wanted the sites to mirror each other as much as possible, while also testing each service's capabilities.
As the commerce solution folks would say, in "just minutes" we discovered we couldn't complete our wish list. As an iCO tech support person explained during a phone call, "On the iCO stores, you're pretty limited. You can't add any pages to it or change any pages. This is a KISS [Keep It Simple, Stupid] principle here, so that new users or uninformed users don't get overloaded with things and try to do too much too fast." In that respect, iCO delivered. We couldn't do too much, and we certainly couldn't do it too fast.

We were able to achieve most of our goals at Yahoo Store, although it required some HTML on our part to create usable content pages and some of the navigational links. Yahoo Store tech support was able to help us with this in a timely fashion. ICO tech support, on the other hand, often consisted of dialing the phone, waiting on hold, and finally hearing a recording asking us to leave our name and number and await a callback.

It's true that Viaweb--which became Yahoo Store in June after being bought by the search giant--has been in business since 1995, while iCO's browser-based store solution has only been available since

April. Arguably, iCO should therefore be given some slack. But then, small businesses looking to build commerce sites usually can't afford to spare the extra rope.

STEP ONE: CHOOSE YOUR WEAPONS

We began our project on July 9, at iCO. After a 10-minute registration process, we started to customize. From four possible store formats, we chose one that included a search field, but abandoned it when this caused section pages--where books were listed--to display entries in an unattractive single column. We switched instead to a two-column layout, then scanned through a list of 67 possible store "styles," offering a variety of images, background colors, and textures-some attractive, some that brought to mind the type of wood paneling found in prefabricated housing. This process alone, which uses Java to display each style individually, can take hours, In the end, we chose the simple Officetek-4, which had an inoffensive white background.

After entering keywords for submission to search engines, we moved on to iCO's Item Entry screen, where we filled in the appropriate information for each book. This procedure was straightforward, albeit unnecessarily slow: At Yahoo Store, you enter product information on a single page. At iCO, you must visit separate pages to load graphics, input product descriptions, include prices, or assign an item to a section.

Uploading graphics--in this case, scans of book covers--went smoothly enough, but iCO's product name field allows only SO characters, fewer than what we needed for some of our book titles. ICO tech support suggested we simply shorten the titles.

Working on iCO was made especially tedious by iCat's servers, which plied us with error messages. We started keeping track of one rather ominous message in particular: "You are an unauthorized user of iCat Commerce Online." By the time we finished building the site, we noted at least 58 of these accusatory notes. Sometimes we could keep working. Other times we were kicked out and had to log on again, only to be faced with a sign-on screen that began, "Welcome to the fastest and easiest way to create and manage a store on the Web."
After a few days of this, we turned our efforts to Yahoo Store, beginning with "Test Drive," a section where you can quickly set up the basics for a store. After 12 minutes, we had filled in company information, and we then spent another 20 minutes making the store official--e.g., giving our credit card number and so forth.

The store editor at Yahoo Store has three interfaces: Simple, Regular, and Advanced. Yahoo Store warns that the advanced editor is meant for programmers: "In Advanced, all protection against damaging your site is turned off. You can easily damage your site beyond repair ...we recommend it only for experienced

programmers." For most of the build, we stuck to Regular.

Unlike the iCO system, which requires you to click a "View Store" button to see changes, Yahoo Store's interface updates changes as they occur. The interface also granted us space for longer titles. We input data for all 11 books under the "Players" section with only a single hitch, a Server Too Busy error. Emboldened with success, we moved to iCO to do the same. After two entries, we were branded as unauthorized and summarily kicked off the system.

The unauthorized user message, an iCO tech-support person explained, "is an issue our tech engineers have been working on. It's in the database. They think they've got it licked."

STEP TWO: MAKE IT YOUR OWN

Much of the building process ran along these lines. Yahoo Store let us upload icons we had created to act as hyperlinks to various sections of the site. We could also customize navigation buttons.

At iCO, you cannot customize the navigation bar; you must use generic, pre-set icons. And you cannot upload images to represent sections-instead, sections appear as text links on the home page. As for our trivia quiz, iCO's software doesn't support content pages. Yahoo Store does, although we had to get help from tech support to find out how to link the trivia answers to the appropriate books for sale.

For our company history and executive message, Yahoo Store includes an "Info" button, which we changed to "About Us," and where we input lengthy information on The Scholar's Bookshelf. ICO's Info page, meanwhile, will display only company name, phone, and address. The only way we could include company history and a message from The Scholar's Bookshelf was to create a Welcome page, something we had hoped to avoid, as it sits between the user and the site contents.

Both iCO and Yahoo Store let you choose payment methods by clicking on which credit cards you will accept. You can also enter shipping charges, which the services will calculate during the order process. ICO grants space for only three levels of shipping charges, however-The Scholar's Bookshelf has seven rates, which vary based on purchase amount, and which we were forced to modify to fit iCO's requirements.

Surprisingly, tech support personnel at neither service knew much about privacy policy issues, although Yahoo Store did let us include check boxes on order forms so that customers can opt out of information-sharing. We were also able to publish our privacy policy on the Yahoo Store site's company information page. You cannot include check boxes on iCO forms; in fact, site visitors are required

to register on-site before making a purchase. Nowhere does the service divulge what is done with collected information.

STEP THREE: BUY, BUY, BUY

With both stores ready-and similar in name and inventory only-we were ready to place test orders. At Yahoo Store, this went smoothly. Shipping charges were calculated correctly; notifications were properly e-mailed to the addresses we entered; and the order appeared on the site and was easily accessed. Yahoo Store will even fax the order to you free of charge.

We tried a similar test at the iCO site, and though we could go through the steps, and the charges were calculated correctly, we couldn't actually place an order. Therefore, we couldn't tell if the system worked and if notification messages would go out correctly. "Part of that," iCO tech support explained, "is because if you do that much, you start filling up your database with a lot of garbage."

Uh-huh.

Excluding the time it took to create images, and including the time spent making corrections and adjustments, it took us about seven full days to complete the Yahoo Store site and four days for iCO. Of course, one reason the iCO store took less time was its relative lack of sophistication. When we pointed out one of the differences--the inability to upload images to represent sections--one iCat sales rep was refreshingly candid.

"Unfortunately, right now iCat Commerce really isn't there yet. It's like a free product, or pretty cheap, so it's a good way to get going into e-commerce" he said, adding that "it's only been out for three months. We're just happy with where we've gotten it so far."

search: COMMERCE, RETAIL

Automated Coding: How the Services Compare

USERS AT BOTH Yahoo Store and iCO can build sites with only a browser, relying on each service's software to do the coding. To see how well each of our Books on Baseball sites was put together, we turned to atWeb's Web Site Garage service,

Each site produced fairly clean HTML, although there were a few errors, said Gautham Godhwani, atWeb CEO and the Garage's Head Mechanic. Neither Yahoo Store nor iCO places height or width attributes on every image, a common mistake that slows down enduser load time. And neither uses .alt tags with all images, meaning users with non-graphical browsers or with graphics turned off would not be able to view page contents.

The biggest problem the Garage found, however, was the lack of

"search engine readiness," particularly at the iCO site, where, according to Godhwani, the keywords and site description used in META tags were nonexistent. This was particularly interesting, as the iCO setup included a box to enter keywords. But in checking the source code for the iCO store's front page, our keywords didn't appear. ICO tech support claimed search engines are still able to see the keywords-an assertion we will look into further in next week's issue.

Neither iCO's nor Yahoo Store's setup process referred to creating the META tag descriptions that replace a site's summary text in search engine listings. We did find it possible to manually insert a script for them at Yahoo Store. This was not possible at iCO. Keywords for the Yahoo Store site, meanwhile appeared in the source code, but the lack of an easily created META tag description was a significant flaw, said Godhwani.
Both services offered good support for a variety of browser versions and platforms. Load times were vastly different-they were much longer at Yahoo-but we attributed this to the fact that Yahoo Store allowed a greater use of graphics on its initial pages.

Although impressed with each site's clean coding, Godhwani said the lack of search engine-readiness shows how far e-commerce services have to go.

"A number of tools out there may be fantastic in getting the site up," he said, "but they don't do much with, 'How do I gain presence and get links?' It will take some time before these tools achieve that."

Andrew Marlatt

A Checklist: What the Services Offer

Site creation

iCat

- Customizable product images

X

- Customizable images within subsites
- Customizable navigation-bar icons
- Opt-in/opt-out boxes available
- Store owner can build content pages
- Automatic thumbnail creation

Commerce capabilities

- Supports secure ordering (via SSL)

X
Automatically calculates tax and shipping
X
E-mails order notification
X
Site tools and utilities
Registers with major search engines
Traffic statistics
Statistical graphs
Tracks overall page views
X
Tracks click-trails of site visitors
Tracks the referring IJRL
Tracks views per page/item
X
Customizable views
X
Help
Online user guide
X
Help articles
X
Technical support
9 a.m.-5 p.m.

PST Mon.-Fri.

Site creation
YAHOO!
Customizable product images
X
Customizable images within subsites
X
Customizable navigation-bar icons
X
Opt-in/opt-out boxes available
X
Store owner can build content pages
X

Automatic thumbnail creation
X
Commerce capabilities
Supports secure ordering (via SSL)
X
Automatically calculates tax and shipping
X
E-mails order notification
X
Site tools and utilities
Registers with major search engines
X
Traffic statistics
Statistical graphs
X
Tracks overall page views
X
Tracks click-trails of site visitors
X
Tracks the referring IJRL
X
Tracks views per page/item
X
Customizable views
X
Help
Online user guide
X
Help articles
X
Technical support
8 a.m.-1 a.m. EST M-F;

8 a.m.-1 p.m. weekends

Same Site, Different

Service

iCat Commerce Online
www.icatmall.com/booksonbaseball

Books:	50
Sections:	5
Pages:	61
Days to create store:	4
Images uploaded:	19
Server errors during build:	58+

iCat Commerce Online
Yahoo Store
www.icatmall.com/booksonbaseball
www.viamall.com/booksonbaseball

Books:	
50	
Sections:	
7	
Pages:	
62	
Days to create store:	
7	
Images uploaded:	
32	
Server errors during build:	
12	

SearcherJan, 1998

News archives: one-stop shopping, boutique hopping and the specialty news search site. (includes related article Battle of the Search Engines)

Author/s: Nora Paul

Those of us who track the availability of news articles through computer access have spent 17-years on a l-o-n-g roller coaster ride. The early 1980s started the slow, clackety-clack build-up of news resources on the commercial services of Dialog. Nexis, and DataTimes. By 1995, availability had climbed slowly to a peak of more than 400 U.S. daily newspapers with all or part of their content available online.

Then the zooming rush to the Internet began with a flood of newspaper Web sites. After a few years of building the sites, the slow, clackety-clack stage started up again as newspaper sites began to figure out how to make their archives available on the Web. Many newspapers see their archives as a potential profit line to throw into the increasingly large black budget hole most sites are digging.

The advantages of Web access to archives for researchers were obvious -- simple searching. low cost. But taking advantage of those factors put searchers back where we started, back on the roller coaster. Instead of having to hop around. calling news libraries to see if they had any background articles on a topic, person, or event, as we did before the rise of commercial online newspaper databases. we now have, to click around the World Wide Web, jumping from site to site hoping to find an archive we could search. The olden days of the large collection of news archives and one-stop shopping on commercial services are waning good news for the searching budgets of libraries, but bad news for the efficiency of searching across publications.

The clear advantage of having access to news sources on the Web through searchable archives appeared in a recent posting to NewsLib (the news librarians' listserv) from a librarian at a small southern

newspaper:

"Hi all. We do not yet have access to Lexis-Nexis, Dialog, etc., in our library yet but plan on (I am pushing for it) getting at least one of them soon. In the interim, however, I'd like to have some access to newspaper or magazine articles online (preferably free). I'm trying to avoid having to go to individual newspaper archives to search. General searches on search engines bring back some good results but often I have to weed through a lot of junk. We've used ERIC a bit too. Any suggestions from others out there who are in the same situation? Thanks."

There is good news for this librarian and others who find themselves in similar situations. Even as the Web's resources expand, it continues to evolve into a more useful tool, as clever and committed people design sites and services to help manage the glut of information. World Wide Web search services which used to brag about the widely flung net they cast out to snag relevant sites have now started to see the sense in selective searching.

Specialty search services.. or as I like to think of them, "shaft search services," drill down into particular types of sites to find relevant pages. Using the same "spider" technology that goes out and catalogs large expanses of the Internet on Infoseek or HotBot or Alta Vista, these robots restrict themselves to specific domains or types of sites. Examples include country "shaft searches" such as Ananzi [http://www. ananzi.com], which just indexes pages from sites bearing the country code ".za" for South Africa. The mailing list search service Liszt [http://www.liszt.com] sends its spider out to locate servers with mailing list software on them and catalogs the mailing lists the software supports.
Other people help manage the vastness f the Web's resources by creating what I call "scout pages" -- pages with links to Web sites that carry the information you need. in narrower and narrower subject categories, people are hunting and gathering, and designing useful links pages. For the news archive Web pages, the "metasite" or "scout pages" approach can help provide a handy list of links and supplemental information about the archives -- but it still requires lots of clicking and hopping on the part of the researcher, because they just link to the individual sites.

Enough specialty search services and scout pages exist to help you find specific, relevant articles from news sites to warrant an overview. The following article looks at options for the researcher who wants to narrow canvassing of Web accessible material to just news-related articles. First, we will look at each of the sites individually, then run a comparison search and tally the pluses and minuses for each.

Searching When You Get There

Before we start, a word about news archives and past issues on Web

sites. Just as news Web sites share no uniformity in design or layout, they also have no consistency in what they define as archives or past issues. Some news sites will hold a week's worth of Web publications on the site. which you must access by date. Others will have several years of an actual archive database available. Some of those story archives are searchable for free as a service of the Web site: others require a subscription to access and search the archive. Some have an archive made up of material that came from the Web publication -- not the print publication (and they don't necessarily define it as such).

Not only is the content and cost of these archives as varied as the publications which provide them, so, too, are the search engines used by the news sites to provide search and retrieval of their available archives. Gone are the days of mastering three search engines (Nexis, Dialog, DataTimes) for the bulk of your news archive searching. Nowadays, each individual news site deals with searching in different ways, depending on the search engine they employed to provide the search.

Some news sites use the search engines of some of the large Internet search services. The Sacramento Bee [http://www.sacbee.com/search/], for example, uses Excite's search engine. Others link to the search engine used by their internal archive. The Miami Herald's library search [http://newslibrary.infi.net/hl/] connects to their Save search software generated database (but, to cloud the picture further, it does not allow all the functions of the full Save search). The Herald is one of a number of Knight-Ridder papers which have their archives searchable through Infinet, a cooperative of newspapers.

Some news sites have started up pretty simple search engines, such as WAIS (one of the early search engine programs designed for Web searching) and made them more search-friendly. A good example is the search service on the San Francisco Examiner and Chronicle Gate Web site [http://www. sfgate.com/wais/search/arch-pro.html]. They have designed query boxes that let you specify desired section names, specific dates, words in the headline, and byline searches. Other databases of past stories on news Web sites work strictly chronologically -- plug in a date and get the Web site pages from that date (but not necessarily the print product's news stories for the same day). To make the situation even more complex, not all news Web sites understand the need for a search manual to help guide searchers in constructing a good query.
Remember, as of now, few of the news Web site archives are as comprehensive as the archives available on the expensive commercial services which usually represent most of the articles that appeared in the print publication. With regards to the news site search services we discuss in this article, remember their limitations. Bottom line, these online resources are good if you need to locate a few representative articles on a topic from major and smaller) newspapers. But do not rely on these news Web search services to

make any kind of definitive statement about how much or whether a particular publication covered a specific story or event.

Scout Pages

Small Hours/News Archives
http://www.aa.net/~rclark/archives.html

List Contents: Rod Clarke, the creator of the list, does not limit himself to daily newspapers. His list includes weekly and monthly newspapers and the archives of news agencies and newsletters from organizations such as Amnesty International or U.N. Daily Highlights and even some television station sites.

List Organization: Organized by regions of the U.S. (East Coast, South, Midwest, West Coast), Canada, Western Europe, Eastern Europe and Russia, Mideast, Africa, South Asia, Asia, Oceania, Central and South America. The links take you to the home page of the newspaper, or, when a site carries an actual database of past news stories, to the search page for the archive.

How to Search: Select a region by clicking on the map graphic and then look at the country listing to find the newspaper you want. (Now, don't you wish you studied harder in Geography class? Try to find a country or state by shape or placement instead of from an alphabetic list of names.) You can also get a screen of the entire list, over 500 sites in all.

The Good News: The list of newspaper links to other countries fills a gap that exists in some of the other "scouting" efforts. Clarke's attempt to identify and link to the specific page which contains archives or back issues of the publication helps since news Web site design is often so murky that it leaves searchers wandering and wondering where the designers stashed the archives.

The Bad News: Some of the links did not work. The listings under individual states or countries are not, for some unfathomable reason, in alphabetical order. What is it about Web site designers? Do they think all the old conventions (like alphabetization -- one o everyone's favorites) must be flouted online?

Suggestion: it would help a lot to have an annotation explaining whether the Web site had a searchable database of past news stories (and the coverage parameters) or if it just provided a news rack with past issues available (where you have to know or guess the date of a particular story to retrieve it). Some indication of the scope of the past information available on the site would also help prevent needless clicking to unproductive sites.

News Hunt http://www.newshunt.com
List Contents: News Hunt is a hybrid -- part scout page, part search

engine facilitator. News Hunt's opening screen says, "THIS IS NOT A NEWS SEARCH! it IS a search of FREE NEWS SEARCHES found on the net!" What it does provide are listings (again, not in alphabetical order or even geographic order) of news publications with free searchable archives.

List Organization: News Hunt organizes around categories of major news needs for archival news sites. The categories then select the archives and display search boxes. The categories comprise General Archives (66 news sites with free, searchable archives dating back more than one week); Arts & Entertainment papers (three alternative newspapers); Breaking News five news services with continuous updates of news stories throughout the day); Business (seven sites from business publications); Popular (three large newspaper sites); Technology (three hi-tech newspaper sections); Weather (two sites). Another section, Newspapers by State, provides a listing of all the publications with archives in each state, a very useful compendium of both mainstream daily publications and niche and university publications.

How To Search: Click on one of the listings given and you get a screen with a search box into which you can type your request. Hit the search button and your request goes to the news site and a results screen appears.

Good News: News Hunt has scouted out archives not just from mainstream news sources but also weeklies and alternative newspapers and business publications. The side boxes next to the search box give good explanatory information about the news site you will search, often including information about how far back the archive goes. Links to the search help screens provided by the news site, if available, aid in putting together a good search.

Bad News: The selection of news sites made available to search is very small. In the category for "Popular" newspaper searches, only three newspapers (The Washington Post 14-day Web site archive, USA Today, and Christian Science Monitor) are available.

Suggestion: Any lists longer than five items should have some sort of organization, preferably alphabetical. Try to find more about the actual scope of the database being searched to facilitate the selection of databases to use.

Kicker: This is a great site for locating and facilitating the searching of news sites with archives, but again, you still end up searching one site at a time.

Special Libraries Association/

Newspaper Division Newspaper Archives on the Web

http://sunsite.unc.edu/slanews/internet/ archives.html

This classic "scout page" was the brainchild of Margot Williams, metro desk researcher at The Washington Post. While teaching an Internet course at George Washington University, she turned her students on to the task of locating daily newspaper Web sites with available archives. The resulting list was augmented with additional links, HTML'd, and placed on the SLA/News Division Web site.
List Contents: Constantly rowing with contributions by news librarians, until recently this list concentrated on U.S. daily newspaper archives, but has expanded to include non-U.S. archives.

List Organization: This long chart, organized alphabetically by state then city (trust librarians!), includes links to the news site's home page, specific links to the archives within the site, database start-up date, and the cost of searching the site (it is amazing to see how many remain free!).

Good News: The clean organization of this scout page makes it a quick and useful reference tool. The information on cost and database start date helps greatly in selecting which resource to choose.

Bad News: Not comprehensive and a bit too selective. but all in all. for a volunteer effort, a very useful resource.

Suggestions: Add more sites from around the world. Also, add alternative and weekly papers. Provide a list of other news search sites (such as the ones in this article) that offer some of that "one stop shopping" everyone seeks.

Meta-Search Pages

These search sites seek to combine in one search what they have found from scanning for news stories from numerous news Web sites. Again, these are good location tools, but you cannot rely on them to give evidence that a certain publication did or did not carry a specific story. They find the stories they find. They are not comprehensive.

Also, these search engines often target the current contents of news Web sites and not comprehensive archives of past news stories. Be careful and aware of what they do. and do not, cover.

News Index http://www.newsindex.com

"News Index, is the original News Only Search Engine. Find breaking news stories and topics of interest from our continuously updating database of articles from hundreds of news sources from around the globe." Begun in April 1996, by the entrepreneurial Sean Peck, this innovative service attempts to cull current articles from news Web

sites worldwide.

Site Contents: A searchable index of stories from "over 200 newspapers and news sources from around the world," including major news sites and Reuters. It indexes current articles only; it is not an archive. In fact, the service supposedly re-indexes target sites every one to two hours. As such, it offers a good way to Let a cross section of the coverage of Contemporary stories, but no way to research old stories.

How to Search: Very simple searching capabilities -- type the words you want in the search box and select either "all" words for an AND search or "any" words for an OR search. I searched for "Hurricane Nora" and found three articles. The trouble with the results was not the fault of News Index, but of the news site that contained the story. Two of the articles (one from U.S. News & World Report, another from the Journal, a northern Virginia news site) had no date attached to the news story retrieved. (Who knows when the story ran?!) The third hit, from MSNBC, displayed the dreaded error message: "HTTP/1.0 404 Object Not Found" (which means the host server can't find the HTML document at the URL you've entered either because it has mistyped the URL, the document no longer exists, or the service you linked from gave it the wrong URL). Results retrieve in relevancy order, based on the number of occurrences of the search terms in the article.

Good News: News Index offers a quick way to review different versions of news events, one of the main reasons the designers give for why people should use the site. In their words, News Index helps because, "Despite what every news outlet wants you to believe, no single outlet delivers the truth, they all deliver their version of the story, and only through reading many different versions of a story, can you finally derive some semblance of what actually occurred." A second reason to use News Index is its ability to provide broader coverage of stories: "Depending on wire services to cover continuing stories certainly means you are not being fully informed. By indexing all papers, you can follow ongoing stories as they happen, not just when a wire service decides something is noteworthy." The third reason they give is economic (and may be the most compelling of all): "The ability to find multiple sources for the same story, if your access level is too low to read a story you want on one service, chances are good with News Index you can find an alternate source and not have to pay to upgrade your subscription." (This will not be good news for those news sites hoping to derive growing revenues from building readership and advertising connections outside their geographic center')

Bad News: In addition to the problems stated above, there is no duplicate detection, so many of the same stories which ran on different news sites appear on the list. The display of the news story abstract in the results list differs greatly depending on the news site the article was retrieved from. And sometimes they do not even

indicate the date the story appeared.

Suggestions: Work harder to verify the accuracy of your links. If you have a problem with some of the news sites removing documents too quickly, stop including results from those sites in the database.

Other Features: News Index -- Delivered is a custom search service which delivers articles which fit the profile you've created to your e-mail box. You can create up to five profile searches that will find only the articles which contain all the words you put into your profile. I did a profile on El Nino, British nanny Louise Woodward, and Cowles Media. Each day, a listing of Web site news stories which pertained to the topics I profiled went to my e-mail box. When I got to work on Monday morning. November 10th, the profile from News Index -- Delivered in my e-mail box contained a listing with links to 21 stories from such sites as CNN, Fox News. The Los Angeles Times, the Las Vegas Sun, and Newsday. Curiously, a notice appeared at the top of the list: "current catalog created at 7:03 on 11/11/97" -- now that's timely! In fact, that's scary. Where's Rod Serling when you need him? I'm not used to getting news delivered a day before it's run!

News Works http://www.newsworks.com
The New Century Network, a partnership of nine of America's largest newspaper publishing companies (including Knight Ridder, Hearst Corp., Tribune Co., New York Times Co.., Washington Post) has put together this site to provide links to the coverage in the more than 125 newspapers in the partners' groups. Each day, the site provides a selection of links to features and news stories from various publications and offers some packages with a list of links to stories about an event or topic from numerous sites. This site showcases the electronic publication efforts from these decidedly traditional newspaper publishing companies.

Why does this site fit into an article about news archives or "shaft" search sites? Because of the feature, found after scrolling halfway down the home page of the site. called Search. The subtitle to Search, "the contents of more than 100 of America's best newspaper is misleading. This search service does not search the contents of these "best newspapers" -- it indexes and searches the contents of these "best newspapers'" Web sites. Big difference! Some of the content on Web pages never appears in the print publication. Much of the print publication's content never gets onto the Web site. Once we're clear on that, we can see what this service does offer the seeker of current news from a variety of sites.

Site Contents: Using Infoseek's [http://www.infoseek.com] spider technology. the spider crawls out to member news sites and creates an index of the contents of the Web pages available on the site. Any pages accessed by entering a password or membership code do not index. (Interesting. At least, one of the nine publishers of New Century Network -- the New York Times Company -- could not include its paper in this national network of newspaper Web sites,

under that restriction. Other member organizations may have papers using registration or passwords.) As with most of these types of sites, they do not make it easy to get a full description of what the spider does and how often it does it.

Test Search: A search for "Hurricane Rick," which had slammed Mexico over the weekend, got eight hits. Four of the eight linked directly to news stories about Hurricane Rick (all of them the same Story from the Associated Press!), the other four linked to a page of the Web site with links to a variety of stories. In two of the pages I could find no reference to any story about Hurricane Rick. I'm sure they had already dropped the story off the page but still left it linked in the index.

How to Search: Use the same search commands and strategy as Infoseek's wonderfully sophisticated search engine allows. Case sensitivity is a particularly useful function for finding relevant news stories with personal or corporate names in them.

Good News: It offers one-stop shopping for many of America's biggest and best newspapers and provides a quick way to get stories run in newspaper Web sites coast to coast.
Bad News: It may be a sign of popularity, but this site loads pages at an annoyingly slow pace. For this reason alone, I would recommend checking some of the other meta-search pages for news sites first if you only want some recent news coverage on a topic. While you can not blame the search service, the fact that so many of the news sites they carry use the same wire stories to cover news events makes for less diversity of coverage than the scope of newspaper sites indexed might suggest. Also. the results page for a search, after you click on a link, takes you to the page, but within a NewsWorks frame. As a result, it often becomes difficult to determine where the story originated or when it ran. Again, you cannot blame NewsWorks: the design flaw stems from individual news Web sites. However, it does seem that the power of this publishing partnership might exert some influence on member papers to see that such information as name and date of publication appear clearly on each page.

TotalNEWS http://www.totalnews.com

TotalNEWS got itself into hot water with news site publishers soon after its launch. Web publishers objected to the frames used on the site which made the retrieved stories stay within the TotalNEWS frame. It looked.. they said, as if TotalNEWS produced the story. According to a New York Law Review article in June 1997, In their complaint. the media companies brought a litany of claims, alleging that the defendants' practice of framing constituted misappropriation, trademark dilution, trademark infringement, copyright infringement, false designations of origin, false representations and false advertising, unfair competition, deceptive acts and practices and tortious interference with contractual

relations." TotalNEWS lost the lawsuit. No longer does the TotalNEWS frame appear around retrieved news stories. Instead it links searchers directly to the news site's full page, frameless, when they click on a link.

This actually constitutes an advantage to searchers. Unlike NewsWorks' search, for example, which keeps the results screen within the NewsWorks page, with TotalNEWS you have no difficulty in determining which site you are on; you see the whole page from the news site.

Site Contents: TotalNEWS' search robot scans about 1,200 news sites and updates the database three or four times a day.

Searching: The system uses a Web crawler called Paradigm News Search. Be sure to read the search help files; you'll see some interesting things, including the fact that searches ignore one or two letter (or numeral) words. So., forget about using TotalNEWS to find reviews in newspapers about the band U2's recent tour or U2 flights over Iran; it's just not searchable!

The Bad News: See above. Also, they provide no clear listing of which sites they do scan.

The Good News: It seems to operate much faster than NewsWorks and constitutes a good backup. TotalNews also seems more savvy about the problem of missing links and the constant flux of news stories on Web sites. Each run of the crawler seems to verify the viability of the database of news story links. Just to make sure that they have the fewest dead links possible, they purge any link older than two weeks. Not good for archives, but the links stay healthier. It's a trade-off.

News-Only Searches from World Wide Web Search Sites

NewsBot (HotBot's News-Only Search Service)
http://www.newsbot.com http://www.wired.com/newbot/

Using HotBot's search engine, this site is all about time. The search box lets you put in a search term and then select from a pull-down box the currency for the items you want retrieved: six hours, 12 hours, 24 hours, 48 hours, four days, seven days, whenever. You can also select certain types of news you want to retrieve: All News, Business, Politics, Entertainment, Health, Sports, Technology, World. Each of the retrieved items has a date and time stamp on it in red, really helping the searcher select the most recent postings. Results display crisply and cleanly

The bad news'? They provide no clear listing of what sites or how often they search. More bad news: They do not purge the database of links to URLs no longer containing the news story. This is a big problem with Web published news search databases. Many Web

publishers re-use URLs for subsequent news stories or simply remove old stories. NewsBot, like it's big brother, HotBot, does, however, provide some of the most innovative ways to focus searches through the search option boxes they offer.

NewsTracker http://nt.excite.com

Excite's news-only search service, NewsTracker provides a number of ways to get the coverage of particular events or specific features from a variety of news sources. A particularly nice innovation on the site is the different display options. Do a search and then have the results sorted by relevancy (with the highest ranked relevancy first) with the lead of the story displayed. Select "Show Titles Only" for a listing of relevancy ordered articles with only the headline and publication name displayed. Click on View by Publication" and you'll see the results grouped by publication name then sorted by relevancy. Or click on "View by Date" and you'll get the results grouped chronologically and within dates by relevancy. This is a great quick scan service of some of the biggest news sites on the Web.

Yahoo! http://headlines.yahoo.com/Full_Cover age/

Get a quick list of top news stories and how papers around the world have covered them with Yahoo's "Full Coverage of News" feature. Click on one of the top stories they follow, e.g., the Nanny murder and reversal of the verdict by the judge, and you'll get a grouping of news stories by News, Magazines & Features, Related Web Sites, Live Coverage, Multimedia coverage. This is another example of the clean, organized, reliable resources put together for Netizens by the good people at Yahoo!. Thank you, Yahoo!.

But we have a bit of bad news to report, too. Use the "Search" box to find stories in the News area of Yahoo! and you get a listing of stories, but no indication in the listing of where the story comes from. You must link to the listing to find out if it was UPI, Reuters, or a press release from Court TV.

NewsHub http://www.newshub.com

Want your news fresh and fresher? News Hub specializes in retrieving only the latest news, refreshing its screen with the latest news every 15 minutes. The older stories, up to yesterday's news, group by time periods, so you can see the very latest but track back to earlier versions.

The Bottom Line

A few years ago users of the World Wide Web lamented the lack of resources to help locate relevant information on the Web. Now, we complain about the glut of choices available to us for different kinds of searches. Just as with the wide variety of search engines available

to users, these specialized news search engines have different strengths, different applications. Get to know them, think about what it is you are trying to find, and select the resource most likely to help you get the job done. And send a note of thanks to those sites that provide good, cheap (even free) access to the resources you need -- they will appreciate the notice.

Software Magazine Sept, 1998

Which way to the Web? (includes related articles on ERP's hidden costs, United Health Services's ERP solution, and on building a new interface)(making the transition to ERP and networked systems) (Industry Trend or Event)

Author/s: Mathew Schwartz

There are a few lessons you ought to know before choosing Build a new front end for your legacy systems.

It's 7:30 PM, you're wrapping up some paperwork when the CEO bursts through the door. "Look," he says to you, breathless, "we've got to get on the Web! I don't care what you do with our legacy systems, but I'm sick of all these green screens." Victim of IBM's recent television advertising campaign? Another example of CEO "drive-by thinking?" Nope, this mandate comes straight from the board: financials, order entry, resource tracking -- get 'em all on the Web, and do it yesterday.

At times like these, you have two clearly defined decisions to make -- buy a brand new, end-to-end suite or adopt a browser paradigm for existing systems. If buying a new ERP suite, then ideally you want it to be built from the ground up for the Web. If, on the other hand, you're happy with your current systems or can't afford the switching cost, you can adopt a browser paradigm without destroying your existing infrastructure -- and your users will be none the wiser.

The Art of the Buy

When Terry Byers walked in the door at American Floral Services (AFS) Inc. in 1996, she inherited a range of ancient systems, and had to decide quickly whether to swap them out or rebuild them. She also had to find a way to satisfy parent company Chemical Bank's request that she look into something called the "Year 2000 problem." Byers decided to scrap everything and install a new ERP package.

The Oklahoma City-based AFS is a "flowers-by-wire" company that competes with FTD and Teleflora, as well as the retail floral industry in general. It works with 25,000 florist customers and also has a large virtual employee base spread across the country, which needs to quickly access customer statements and profiling information.

Byers took stock of the existing situation. Founded 20 years ago as a mom-and-pop business, AFS was still using an antiquated character-based HP legacy 3000 application, as well as some smaller AS/400 systems like Apex and older accounting software from Great Plains Software and Peachtree Software. There was no network in sight. All in all, not a lot to build upon. "Our legacy application, of course, was not [Y2K-compliant] and making it so would have been expensive and time-consuming," says Byers. On top of that, she didn't want to just have a working 20-year-old legacy application by the time she was done.

Management, having brought her in, didn't need to be sold on the idea. "The company was starving for it. We were absolutely starving," says Byers. "The typical data was in concrete, you had to get programmers to write reports, it was just a ball of spaghetti code."

Byers looked at about 20 ERP packages, primarily from a functional perspective, and quickly zeroed in on packages from Lawson Software and Oracle Corp. Lawson was already Web-enabled, while "Oracle was selling futures but was supposedly very near," says Byers. The ability to operate through a browser was essential for AFS, because of the immense numbers of people to support. Also, when comparing the two, she says she wasn't as comfortable with Oracle's service level. Furthermore, Lawson would warranty its product as being Year 2000-compliant. So after referencing 15 of Lawson's customers in various stages of implementation, "I felt we really knew how it would go," says Byers.

AFS selected Lawson Software, a Minneapolis-based $166 million ERP software provider with an impressive 97% customer retention rate. Currently, AFS has Lawson financial software installed, and they built their own sales force automation tool. For mobile users, "one of the things we found is that Lawson runs so thin we were able to install Lawson on our laptops -- [in-house-users] are not running through a browser," despite the fact that they are able to remotely access 1,000 data tables in an Oracle database, says Byers. The client-side OS is NT, and on the server side it's HP Unix.

Don't let Lawson's size fool you: "They compete in the enterprise applications market with companies like SAP, PeopleSoft, Oracle, and J.D. Edwards. They hold their own among very notable players," says Clare Gillan, an analyst with International Data Corp. in Framingham, Mass. "They've had consistently strong growth for a number of years," she says, particularly in the vertical markets they target-- healthcare and retail.

The latest versions of Lawson's software are all Web-enabled; the company rebuilt all of its applications with object-based architectures and Web compatibility. This functionality squares with Gillan's advice for companies looking at ERP: "If you want an entire electronic commerce solution, look at a tool optimized for the Web." Objects in Lawson's architecture can be assembled for various kinds of front ends, browsers included. In addition, Lawson created a "self-evident front end" that gets kudos from users for its smart GUI.

Running a Restaurant

Financial planning and sales force management have devotees outside the floral industry. Rainforest Cafe Inc. needed an enterprise solution, especially where financials were concerned. The $108 million company, which won the 1997 National Retail Federation's Small Store Retailer of the Year award, started in Minneapolis in 1994, and is still headquartered there. You might have encountered its exotic, jungle-theme restaurants and retail spaces in malls. So far, there are over 20 Rainforest Cafes across the country, including one in Disney's Animal Kingdom, as well as several abroad, with many more to open soon.

Rainforest's revenues between 1996 and 1997 grew a whopping 260%, and the small-scale systems in place weren't giving management the kind of information it required, or the ability to forecast company performance or needs. The company knew it wanted an ERP package; it was just a question of which one. The overriding goal of the installation, says Mark Rabinow, Rainforest's CFO and also me person who spearheaded the ERP adoption effort, was "to have our entire company be on a common database -- we had to maintain a backbone to the World Wide Web at a minimum expense." And as a newer company in the volatile restaurant industry, they didn't have time to waste when implementing it.

After evaluating various ERP packages, Rainforest, like AFS, settled upon Lawson. "Lawson had the Internet connection and functionality in place, really nobody else had that at that time. We also just thought the software was better value than other people we looked at, like PeopleSoft and SQL Financials," says Rabinow.
Though Lawson had originally budgeted about 12 weeks for the summer 1997 project, "it was a quick implementation -- about 6 weeks," says Rabinow. Currently, Rainforest uses Lawson Insight Financials for worldwide accounts payable and purchasing, as well as to provide realtime financial reporting to about 40 users -- 20 in the home office, 20 in the field.

Implementation speed is one of the cornerstones of the mid-range ERP sale, and all in all, Rabinow says implementation at Rainforest went smoothly. "Certainly, we had some protocol issues to work out with the Internet, which Lawson got resolved pretty quickly," he says. "Was it perfect? No. Was I satisfied? Yes."

When senior management is so directly involved in the implementation, the old business/IT cultural disconnect issues are definitely blunted. At American Floral Services, management quickly signed off on Byers' ERP project. But once installed, she realized that she still had a lot of political work ahead. Many business managers think that installing an ERP system will allow them to lay off IT staff who previously supported the legacy systems. Managers at AFS said, "We've put in the system, how many heads can we cut?" It's not about heads rolling, Byers would say, it's about moving resources from managing a project to supporting it, or to other projects.

"The thing to remember with business-initiated changes is that expectations run very, very high," she says. How do you educate management? "You can't say it once or twice, you have to be in rewind mode," says Byers. Eventually, management realizes that "it just takes time" to implement these kinds of packages. Changes don't magically appear with a flip of the switch. Improved productivity doesn't happen overnight -- since in stalling Lawson, says Byers, "we're still looking to gain some of the productivity, though we've seen some."

Sometimes, expectations manifest in ways IS didn't foresee. Take report writing. Although the Lawson software has "incredible drilldown," management still likes the old-school paper reports, she says.

But when it comes to finances, there aren't any surprises. In an age where every dollar spent on big ERP software often necessitates $10 to $20 in further "hidden" consultant costs to get it working, Rainforest CFO Rabinow notes that Lawson didn't try to pull any fast ones on them. When it came to budgeting, "they were actually very accurate," he says. "Whereas some software companies lowball you and then the consulting time runs amok and you end up spending another hundred grand more than you thought. Not the case here. They didn't try." So in the end, Rabinow figures that Lawson was still "within a couple percent of what we had originally projected and agreed upon." Furthermore, as long as a company is on a maintenance contract with Lawson, upgrades are free.

Because of Lawson's object-based design, both companies are continuing to introduce features since going live. AFS is continuing to roll out SFA functionality every three to six months. After the financial suite, Rainforest added HR, and still plans to implement the performance indicator suite, also a Web-based feature, to further track store profitability and ROI. In addition, Rainforest will manage capital expenditures for new store construction. At $6 to $12 million per project, and with a number of projects currently in progress, having realtime data directly affects Rainforest's ability to stay lean.

Live to Fight Another Day

It's the gadget factor -- given the choice, most IS managers would probably prefer installing a new ERP system instead of knitting their

legacy systems together with a new GUI to get them to the Web. But many CIOs faced with Year 2000 problems are thinking maybe now isn't the best time to fight the enterprise packaged application battle. Or maybe your firm just doesn't have the money to spend on an ERP package.

In fact, many organizations, especially those outside the Fortune 500 realm, often have more modest needs than a new ERP package. "There are existing systems, and there are ERP systems, and a lot of people are sticking with their existing systems and then expanding them over time," says Peter Kastner, an analyst at The Aberdeen Group in Boston. Even after installing an ERP package, most mainframes still don't end up on the street curb. "State Street Bank," notes Kastner, "put in SAP, but they still have mainframes, big time, because they still have huge trust accounts." He predicts that bath Unix-based ERP packages and mainframe legacy systems will be running side by side for years to come.

Why not migrate toward a multi-tier setup with PCs and NCs on the desktop, doing it in smaller steps, swapping out older hardware for new, and taking advantage of the ease of training on browsers, as opposed to many client/ server systems? That's the solution that Universal Health Services (UHS) Inc. in King of Prussia, Pa., decided to pursue. The third-largest for-profit hospital chain in the U.S., UHS has over 6,000 beds across 25 states and Puerto Rico, and had 1997 revenues of $2.9 billion.

UHS was saddled with a legacy system that had to be upgraded if the company hoped to remain competitive. "It's only in the last few years that hospitals have viewed themselves as competitive," and have become more interested in investing in technology to hone that competitiveness, says UHS CIO Linda Reino.

Functionality doesn't have to come at the expense of huge infrastructure upgrades. Now that "legacy" technologies have come full circle in the IS community, newer technologies are making it increasingly easy to unite the mainframe with the Web. Previously, companies risked losing one of the benefits of mainframes -- centralization. "One of the most important things from my perspective is the ability to have as centralized an infrastructure as possible," says Reino. When keeping patient records, for instance, supporting a Unix database in every hospital is overkill if the current mainframe system can handle all of them.

And, of course, now, "the mainframe is acceptable because it's a big server," says Reino. Call it mainframe retro. One of the consequences of mainframes suddenly being classed as enterprise servers, data repositories, and warehouses is that suddenly "we can't throw out the investment," she adds.

When it came to enhancing that investment, Reino hadn't seen a solution that fit her needs, which include increasing the quality of

patient care and physician access to data. "The physician is a big customer," she says. GUI packages didn't suit her needs, because they just "sit on top of the mainframe -- old PF [program function] type methodology," she says. But then she evaluated Opal, from Computer Associates, and was sufficiently impressed. "Now you look at a product like Opal and say, `Wait a minute, now I have an opportunity to actually interact with the menu navigation system that I have in place right now. And I can alter it and make it more streamlined for my users so not only am I sitting on top of it and making it look more attractive, users will think they're using something state of the art,'" says Reino. By condensing screens, people can also do their jobs faster, and by moving to a browser paradigm, training time is dramatically reduced. Why should they care if there's a mainframe back there?

The urge to repurpose existing systems is clear, says Robert Lincoln, Opal marketing manager: "In a buy vs. build, a lot of times the issue has been rebuild vs. scrap." Companies with heavy investments in character-based applications aren't always going to want to quickly scrap them for client/server systems. "The image you tend to get is that everyone is running wholesale toward E-commerce at breakneck speed, but when we get to the customers, they're doing it slowly," says Lincoln.

Aberdeen's Kastner says that Opal fills an underserved role by putting "a modern, lightweight front end on literally hundreds of thousands of dumb applications that are out there." And by moving from dumb terminals to the browser, "I've just torn out the whole network without sacrificing any of the security that I care to have," he says.

Let's play devil's advocate: Isn't Opal mere screen-scraping technology? To an extent, yes-- it reads and writes to terminal screens by tricking the database into thinking Opal is just another CICS session. But, notes Kastner, "what they possess is more than screen scraping, because you can make the applications more graphical, use the look and feel of the browser," and trim menu hierarchies down to size by importing only needed information.

Opal creates object representations of the internal workings or a database in effect, a component environment filled with encapsulated objects and processes. "So in a drag-and-drop way, you can link a table to a corresponding object on a host screen," says Lincoln, creating a codeless, easy-to-use development environment, to the point where users don't have to be hardcore programmers. One university uses student interns, for instance, to do much of the work. "A lot of customers for Opal don't have their own development staffs anymore," says Lincoln. Besides screen scraping, or providing host-to-Web style emulation -- from 3270, 5250, VTxxx, HP 700 -- Opal can monitor the connection and state of the session, which HTML can't do, because it's stateless, as well as provide desktop deployment. But more impressive is Opal's ability to integrate information from different databases Ingres, Sybase,

DB2, Informix, Oracle, SQL Server (see sidebar, "Building a New Interface," pg. 76).

At UHS, Reino is clear about her goal for the Opal implementation, currently in progress. "Opal is going to simplify our existing workflow, reduce the training and maintenance burden on our help desk, and increase the functionality that's available to our users -- and that's what it's all about."

Begun a year ago, results are already apparent. "We've already dramatically enhanced the clinical capabilities," says Reino. Many hospitals now have online access to the POET (patient order entry tracking) system. Information pathways currently available to physicians in 3270 format are being recoded into Opal. All in all, about 75% of the target functionality is in place, with a full-fledged pilot to begin soon at Manatee Memorial Hospital near Tampa, Fla. After a more extensive pilot in 5 to 6 hospitals on the UHS technology taskforce, "we think by next year we'll have this up and running in our physician offices," says Reino, and then speed the rollout to the rest of the hospitals. "I want to get the same pathways into the hospitals, and to do that there might be some wiring and infrastructure changes that I have to do in the facility to give access," she says.
With an organization of UHS' size, using Opal is as strategic as it is functional -- with an overall price tag to match. When she talks to other senior managers, "I don't sugarcoat it," says Reino. Although investing over a current infrastructure always provokes questions, "the message was I want to make this investment and keep this in place vs. what I think is a more significant investment." Cost was definitely an issue, as were security concerns and leveraging previous mainframe investments.

But UHS should gain system longevity: CA claims that Opal is part of a larger "Harmony" strategy -- a group of applications that share a common framework, so any investments a company makes today for short-term results, say 12 months, will not be entirely tied up in a certain kind of technology in effect, they shouldn't become "legacy" systems right after you buy them or over the next 3 to 5 years. As everyone in IT knows, "interim" solutions redefine what the word "interim" really means.

Opal is relatively new, and UHS one of the earlier health-care customers. "The only letdown I've had is the amount of education that I've needed to provide to the CA folks," says Reino. "I don't mean that from a technical standpoint, this is from the health-care industry standpoint -- they did not have a lot of strong health-care industry skills, so I'm trying to sensitize them to the needs of healthcare. And I've been very happy with the way they've responded." Nevertheless, "there's a good piece of my hide riding on the line," she notes. "CA needs to work with me to let me know when there's technical roadblocks and infrastructure roadblocks --

it's a learning experience, on a daily basis."

The Land of Cheap Labor

On the surface, the goals of the University of Miami (UM) were not entirely different than those of United Health Services. UM was happy with its current IMS mainframe database and had spent a good deal of time customizing it and other systems to provide student access. For instance, in 1992, Michael Zucker, UM's assistant director of application development, and Beth Hernandez, senior systems analyst, created a student system called Easy-- as in, "easy to use." By disseminating administrative data, it allowed students to check their own information, note when employers were coming on campus to conduct interviews, and keep abreast of other day-today information. "Beth and I were very excited, we created a whole student system, great and easy to use," says Zucker.

Au contraire-- once in use, "the students said no, it's difficult and cumbersome to use," says Zucker. Typical 3270 mainframe problems -- numbers in letter fields locking the session -- were partly at fault, and different IS and student expectations probably factored in somewhere too.

After evaluating Opal, UM became an early Opal beta customer in late 1995. "If you have legacy systems -- something that works -- and it does everything you want except the interface, I think it's crazy to throw it away and spend millions" just getting a GUI, says Zucker. With a PeopleSoft installation costing $5 million, he says, justifying that expenditure in UM's case is impossible. So the primary reason for using a product like Opal, Zucker says, is to keep the legacy system. "In our case Opal products, or products like it, are an absolute necessity. We can put information on the Web in student labs without throwing away everything we've worked towards for the last fifteen years," he says. The secondary reason for choosing Opal is because developing on it is "an amazingly cheap cost," he says. Pricing starts at about $1,000 per seat.
For development efforts, the IS department tapped engineering undergraduates. "University," Zucker notes, "is like a prison -- lots of captive, cheap labor." No doubt students entering the workforce might spin it differently. Regardless, they're a group ideally suited to the 5 to 6 month IMS learning curve. And graduates train their replacements.

The initial implementation was quick. "Working with a team of students, Beth was able to knock out a very large system in six months" that contained all the core functionality the university students needed, says Zucker. The hardware investment was two boxes -- one to run the Web server, the other the Opal server, which is also connected to the mainframe via TCP/IP in a terminal session-like mode.

Now, Easy is available in all the student labs, which mostly have

PCs, and via every terminal.

That's just the beginning. "We've been adding to it since then, enhancing it," he says. 12,000 students have to interact with the system heavily. It's the only place to vote for student elections, for instance. Or to register for classes.

Not only that, but they can do it over the Web. This is important for the increasing numbers of students who dial in from off-campus. Because of a plug-in on the client side, Opal is essentially client/server -- just over the Web and through the browser, and because it doesn't use HTML, Opal doesn't have any of HTML's graphical limitations. The plug-in also authenticates users when they log on to Easy.

The University of Miami had fewer hardware upgrades to deal with than UHS. It invested in a new Dell server, strictly for project tracking. The ability of Opal to front various applications is also finding use at UM, which runs Ingres, Sybase, and DB2, among other things. Opal is "a nice way to integrate other relational data with legacy data," says Zucker. For the past 10 years, the IS department has tracked detail history, "but it's something we very seldom looked at." Though it's more of a minor application, the server now lets them investigate problems on an ongoing basis.

UM has been pleased with the results. "If you have a legacy application that's fulfilling 90% of the business need and you're just worried about presentation and appearance, this is a great solution," Zucker says.

UM currently has no plans to surrender its mainframe systems, but rather is thinking about how to extend them. Luckily, Year 2000 code remediation was begun in the early 1980s; mainframes are there to stay. "Maybe in ten to fifteen years," says Zucker, "we'll be the last person running mainframe data." Then again, maybe mainframes will be back in vogue.

RELATED ARTICLE: ERP's Hidden Costs

Why is it important to focus on the Quality of the user interface and Browser access to canned reports when you're installing an ERP system? Because if you don't, you're going to find huge cost increases when it comes time to install the system and generate reports.

In a study Lawson Software conducted with Price Waterhouse a few years ago examining how consultancies perform ERP integrations, they found that"30% of the money you're going to spend with them is writing reports that are basically there to extract key performance information for senior management," says Mark Galloway, Lawson's vice president of strategic planning. Curious-- you'd think system integrators Would have more ready-made templates in place. With Lawson, that's "out-of-the-box capability," says Galloway. Once the

customer describes their business by setting up a general ledger sheet, it gets key performance information-- in a variety of report formats-- ready to go.

RELATED ARTICLE: Behind the Curtain

Mainframes are great-- for some things. For instance United Health Services' old mainframe is still an excellent data repository for patient records. But you wouldn't want to try and store performance information or security tracking in it; it just wasn't designed for that.

Using Computer Associates' Opal, the King of Prussia, Pa.-based UHS, which runs a chain of hospitals, references an additional accessory database, since the mainframe can't really provide a streamlined audit capability for the hospital. But the nice thing with Opal, says UHS CIO Linda Reino, is that the user doesn't know and doesn't care how the information is being assembled. As long as there's ease of use, they're happy.

UHS is also using Opal to front a database that maintains an audit trail of which charts doctors look at. Not only can they tell if doctors are properly checking all of their charts, but if they're looking at unauthorized charts, then the hospital is aware. Maybe one doctor is covering for another -- then again, maybe not.

Such security is no small concern. As a health-care provider, UHS has very specific parameters set by the Joint Commission for Accreditation of Hospital Organizations (JCAHO). If JCAHO doesn't like a hospital's security, then it won't accredit it, which means no one else will deal with the hospital. "I have this unique dilemma where, on one side of the fence, I have my security requirement and my regulatory agency saying you cannot provide free access to data," says Reino. Then, on the other side, there are the users, doctors "Look, if you make it hard for Dr. Brown to have access to information, he's going to piss and moan," she says. With Opal, she hopes to provide "reasonable access with appropriate security parameters in place."

RELATED ARTICLE: Building A New Interface

Screen-scraping technology in Opal is high on the University of Miami (UM) list of successful features. "That's one of the advantages of Opal--what was the fancy term? Repurposing. We were able to take some of the administrative screens that we hadn't made student versions of... let's say they're 50 fields, you just scrape three off," says Michael Zucker, UM's assistant director of application development. Even if a packet sniffer is used during the session, students still don't have access to the unrepresented parts of the screen--the information just isn't transmitted out of the mainframe for student sessions.

The ability to condense screens was an integral part of United Health Services (UHS) choosing Opal as well. "Instead of a user having to

go through seven, eight, nine, ten screens from a revisionist standpoint, we have the ability, using Opal, to combine screens," says UHS CIO Linda Reino.

Combining data from disparate screens simplifies workflow, secures access to different kinds of information, and allows for a new GUI--"one of the big d rivers is that I can sex it up a bit," says Reino of the interface. "Instead of saying the words onscreen that a woman has had a migraine, she says, a screen pops up with a graphic of a woman. The trouble area is highlighted, and the symptom listed next to it. It's a boon for doctors who'd rather process graphic, as well as textual, information.

On the other hand, the University of Miami is shying away from making its student system, Easy, too graphical. "We've talked about that and argued internally," says Zucker. At issue: How much graphics really slow down a Web session, compared with how much they can enhance it. "Our compromise is that for pages that are like a menu we'll throw a picture on, but the ones that are the registration screen, we leave the pictures off," he says. This is because lots of students don't use broadband--many still dial in on 14.4 modems.

SearcherFeb, 1998

Wins and losses: the latest developments in online competitor research.

Author/s: Helene Kassler

Marriages, wars, births, deaths, mergers, acquisitions -- a Shakespearean tragedy? A Wagnerian opera? No, just 1997 and another year of tsunami-scale changes in the online world with impacts touching many areas of competitor intelligence (CI) research.

Last year's changes clearly spotlight a transformation in the online industry. I -- and everyone else -- have seen the future and it is the Internet. At Fuld & Co., a consulting firm specializing in competitor intelligence, we find ourselves increasingly relying on information gleaned from both free and fee-based Internet services. One example of a free resource: A client retained us to determine the technological base for a competitor's electronic commerce Web site. My search on Monsterboard, the commercial job Web site (free to searchers), displayed more than 100 job positions open at that firm -- revealing vast details about the competitor's hardware and software use.

At my company, we even pay for Internet-based information services, when they offer value. Several project managers at Fuld subscribe to an affordable Internet-based, direct-to-desktop alerting service. Some profiles retrieve industry headlines; others specify company or product names. All the service's users find value in news, filtered to their needs, delivered to their desktop in a timely fashion.

It seems clear that the most exciting and useful services and changes are now developing on the Internet. While the older, mature companies port their massive databases over to the Internet with few Internet-enhanced innovations, in contrast, companies building new services from the Internet ground up (usually newer, smaller firms) often offer greater functionality by taking full advantage of the Internet's innate abilities such as hyperlinking,

graphics, and interactivity. Even within the Internet world, the rate of development and innovation is staggering. The game of competitive leapfrog takes place in a matter of months. One Internet moment finds a particular service the darling; three months later, a brand new service emerges with advanced features, greater content, a simpler interface, or a lower cost.

While thrilling, the Internet world can also be challenging and exasperating. Thrilling in its new services, capabilities, and usually low costs; challenging to keep up with new resources; and exasperating when one finds a favorite Web site redesigned with a favorite function buried five mouse clicks deep, or worse the site has vanished, or even -- heaven forbid -- now costs real money!

Fresh New Faces

In the commercial online world, innovations offering the greatest value to CI researchers have arisen from creative new systems or upgraded services from companies such as IAC, Manning & Napier, and Northern Light. Among the big three traditionals (Dialog, LEXIS-NEXIS, Dow Jones), only the new Dow Jones Interactive service utilizes the Internet's native talents at all well, and even there, restricts its "Netted" offerings to a special section featuring business Web sites. The Web's natural functions have not yet been completely integrated into the actual Dow Jones databases. Currently, hotlinks in the Wall Street Journal Interactive Edition articles generally take you to Dow Jones' own Business Briefing reports for more background information on companies -- material included in the $49 annual fee. Sometimes it refers to an article in the Publications Library for which Dow Jones does charge extra. If a reporter links to an outside Web site, the WSJ Interactive Edition will put the URL in parentheses in the text of the article.

In contrast, IAC's InSite Pro Web-based service, introduced in 1997, offers tremendous assistance to CI research through its use of hyperlinks. Top IAC databases (including PROMT, Trade & Industry, Newsletter, Computer, and Company Intelligence databases) are searchable in easy-to-use, simple, or advanced interfaces. But the features of greatest value appear in the resulting full-text articles. Running along the left side of the screen, IAC has included hyperlinks to other stories gathered and categorized by IAC's vast index fields. For example, one article retrieved in a search on Gillette and antiperspirants displayed hyperlinks to stories about the named advertising agency, as well as to the antiperspirant and advertising industries. In addition, when available, IAC includes hyperlinks to corporate Web sites in the Company Intelligence Database (currently 9,000 links). The company plans to include corporate hyperlinks in the remaining databases in 1998. By harnessing the Internet's innate hyperlink function, InSite Pro makes intelligent connections, conveniently extracting and offering many of the resources we, as CI experts, would pursue in additional new searches. With its flat-fee pricing, InSite Pro allows you to follow those hyperlinks to your

heart's content.

Manning & Napier (http://www. mnis.net) is another new entrant with a flat-fee, Internet-based service called DR-Link. The service offers several useful features, particularly its natural language search capabilities and assistance in focusing a search. In 1997, the service introduced a new "Visualizer" Java function that can prove to be a great asset in CI research. In response to a query about biotechnology research in vaccines, for example, the system produced a hyperlinked chart, visually displaying and separating out companies and organizations, people, countries, and subject areas related to the search -- all valuable facets for a typical CI project.

Northern Light (http://www.northern light.com) represents one of those fresh new Internet faces mentioned above, in this case, a company with a novel idea for a commercial online service -- a highly accurate yet simple search engine offering inexpensive full-text articles from more than 400 business and industry publications combined with targeted Web searching. Based on a mass marketing model (sales in the millions of transactions to the general public as well as professional researchers), Northern Light's pricing structure charges per article, with no sign-up or online searching fees. Northern Light also offers features that ease the CI research process. Search results, automatically ranked for relevancy, include a list of custom search folders which hyperlink to stories categorized by subject, industry, publication, etc. Again, these hyperlinks help focus quickly on results, eliminating the need for new searches. Northern Light's excellent search engine can also search the Web with highly accurate results.

A new Web site called CEO Express (http://www.ceoexpress.com) is a business researcher's dream. Completely revamped this year for more intuitive use, the site features one-stop shopping with a vast, well-organized collection of essential company- and business-related sites and services on a single page. Aimed at "busy executives," CEO Express not only links to the more obvious news, government, and business sites, but also connects to overlooked sites such as Chambers of Commerce, the Internet Bankruptcy Library, and software download sites.

Last year, the online industry news story with the greatest coverage was clearly M.A.I.D. PLC's purchase of Dialog from Knight-Ridder. It's too soon for the impact from the sale to reach the searcher. Latest reports promise a marriage of Profound's intuitive Web-based interface with Dialog's vast databases in the future. However, in 1997, the big push was DialogWeb. This product offers little of inherent value to information professionals performing CI research.

However, Dialog's addition of company index fields to close to 80 files (including IAC Newsletter, Early Edition, McGraw-Hill, and American Banker, among others) has proven to be a great help, allowing index searching for company names in files whose producers did not see fit to include this basic descriptor field. Dialog-

generated company names have now gone onto almost every news source added since January 1997.

Dow Jones Interactive, whose Publications Library at presstime extended to over 5,000 titles and 70 million articles, also instituted a vendor-based company indexing system built on sophisticated software. The indexing provides standard company name and ticker symbol access across the whole range of news sources, regardless of whether the database producers provide their own company name descriptors.

In a "marriage of convenience," Dow Jones, the Financial Times, and Dialog allied to produce the World Reporter. When the World Reporter database reaches its full strength (currently it includes over 400 publications), it should prove worthwhile for international CI research, offering a company index field, news from 500 major international publications in 26 languages (currently 17); and translated abstracts or full-text articles in English.

The Net's Database Evolution

With most major companies supporting home pages, the Internet has grown into a vital information resource unto itself -- not just an aggregator or alternative platform for commercially available databases. The Internet now serves as a data source itself, combined with a searchable advertising medium. Although companies clearly hype their products, image, stock, etc., they indeed provide valid and valued information through their Web sites. Useful CI information such as news releases, SEC filings, job postings, executive biographies, conference participation, speeches, all commonly appear on company home pages. Moreover, even "hype" can have value -- knowing how your competitor portrays itself, its products and services is fundamental to understanding your competition.

At Fuld, we typically begin research projects with a thorough review of a company's home page. You can often guess the home page for a company (http://www.companyname.com). Garden variety search engines often have difficulty zeroing in on a large company or a firm with a common name (imagine looking up Apple's home page). So when the shortcut try doesn't work, we use several company look-up Web sites. Fortunately several new services have arisen to replace the Domain Name Lookup Service, which managed to redesign itself into inadequacy last year (http://www.internet.org). Several new favorites include NetPartners Company Locator, which uses Internic registration information, thus offering the most extensive database (http://www.netpart.com/resource/ search.html). Several other sites focus on larger, predominantly public companies: Hoover's (http://www.hoovers.com); CompanyLink (http://www.company link.com); Companies Online (http://www.companiesonline.com); and the multi-function, ubiquitous Yahoo! (http://www.yahoo.com/Business/Companies).

Drilling Down Locally

Local press, always a fertile field of useful information, offers intensive and extensive coverage of hometown companies. Revealing stories on company activities as major as reorganizations or as minor as OSHA inspections will receive more coverage in the local press than elsewhere. Dow Jones continues to be the leader in this arena, expanding full-text coverage to more than 500 newspapers, both large and small. However, the Web-based Dow Jones Interactive service eliminated a useful feature from its news retrieval system -- the ability to scroll through a list and limit your search to a whole state or a single newspaper. Instead, with the Web-based service, you must know and use obscure keyboard commands (such as sn=boston globe) to search individual newspapers. However, according to Dow Jones, the company plans to bring this functionality back to the Web product next year. On the other hand, only the Dow Jones Interactive product features relevancy ranking and a free peek at the first sentence or so for each article.

Dialog also continued its increase of local press coverage to more than 100 local papers. Weak spots such as Texas, which had lacked current news in the Dialog collection, now have local press coverage.

Sadly, 1997 saw the passing of a local press pioneer. DataTimes, an early advocate of local press offerings, ceased separate operation and was integrated into UMI's Proquest Direct product. Under an agreement with Dow Jones, UMI will direct DataTimes customers to Dow Jones Interactive, which will pick up local press formerly offered by DataTimes.

LEXIS-NEXIS, the pioneer of full-text news online, added Bloomberg News in 1997 with corporate news coverage from around the world, along with extensive reporting on governments and financial markets. They have over 60 of the nation's top 100 newspapers available full text, with 15 local papers added just in 1997, along with lots of wires and business news sources. LEXIS-NEXIS appears to have a steady growth policy when it comes to U.S. and international news sources.

Hometown papers also have a strong showing on the Internet. For one client, we sought information about a pilot energy project in several small New Hampshire towns. Precious little appeared on the conventional online services; even Dow Jones at that time had few full-text offerings from New Hampshire. Instead, we turned to the Internet and ultimately found crucial information, including invaluable contact names, through small local newspapers on the Web.

Newsworks (http://www.newsworks. com), the first site offered by New Century Network, is a joint venture of nine publishers featuring 125 newspapers as well known as The New York Times and as obscure as the Anchorage Daily News. The beauty of the site is that

you can search all papers through one interface with one syntax and retrieve free full-text articles. The beast of the site is that you still never know if you'll retrieve a full-text article or an outdated dead link to a non-archived article.

If your situation requires that you know full-text costs beforehand, you can head over to Newspaper Archives on the Web sponsored by the Special Libraries Association News Division (http://sun site.unc.edu/slanews/internet/archives.ht ml). This site features a chart with per article charges plus hyperlinks to full-text archived newspapers. For those with a subscription to BiblioData's resource-filled CyberSkeptic's Guide to Internet Research or Fulltext Sources Online, you also have access to BiblioData's "Private Zone" (http://www.bibliodata.com/private.html), a password-protected collection of more than 400 newspaper and magazine sites with archives both free and full-text.

American City Business Journals (http://www.amcity.com), a site noted for local and regional business news, features 35 regional business papers. A major revision this year increased its utility. This is one of the few sites offering free full-text archives and a single searchable interface. Sometimes the Internet still offers a free full-text cake and lets you eat it with a single searchable interface!

Clipping Without Scissors

Electronic alerting and clipping services can also prove very useful to CI researchers. It's hard to argue with an e-mail alert about a competitor waiting in your mailbox when you arrive at work in the morning. This year has seen several changes among Internet-based alerting services. The most promising change is the extensive service upgrade of Farcast, which also renamed itself INQUISIT. (http://www.inquisit.com). Once limited to wire services, INQUISIT now serves up inexpensive Internet-based alerts and industry "broadcasts" derived from 400 newspapers and periodicals around the world. Simplified templates help set up a personalized alert on a company, product, or name. Headlines, paragraphs, or full text appear directly in your e-mail box or are retrievable via the Web. The INQUISIT broadcast service delivers a list of industry-specific headlines several times a day, allowing you to check off the stories you want, which they then deliver to your mailbox, full text, within 20 minutes.

The creation of NewsEDGE (http://www.newsedge.com), formed by Desktop Data's merger (dare we say acquisition?) with Individual, Inc., prompts many questions. While Desktop Data's high-priced offerings focused on real-time, corporate-wide news delivery, Individual's claim-to-fame centered on inexpensive or even free personalized news. The newborn NewsEDGE will have more than 2,000 content sources under license. The merger may be good for the companies involved, but the endurance of Individual's affordable services remains to be seen. To quo Han Solo, "I don't have a good feeling about this." After all, how will we, as customers, benefit from

the high-priced spread merging with its low-cost competitor?

A new wrinkle in alerting services can help researchers monitor competitor Web sites. Offline Web browsers that download full pages from a Web site are predominantly designed to peruse Web pages offline. However, some will also send alerts when specified Web pages (including keywords or passages) change. You can set up some offline browsers to monitor a competitor's Web site for changes on pages holding SEC filings, news stories, job postings, new products, management changes, executive speeches -- all keys to competitors' strategies. Two commercial packages of note are Tierra Highlights2 from Tierra Communications (http://www.tierra.com) and Web-Whacker from the Forefront Group (http://www.ffg.com). Both software publishers upgraded their products last year. For lists of other offline browsers, visit Tucows (http://www.tucows.com/us. html), a well-regarded gathering place of Internet software. When prompted, identify the continent, country, state of your location, and operating system. Then choose Offline Browsers from the chart of available software.

Patently Clear

Patent searching is critical to an effective CI research program. Patents point to new areas of R&D, to products slated for future commercialization, and to inventors/experts useful as contacts. Several Internet-based innovations now make patent information more accessible to a wider audience. Introduced early in 1997, IBM's Patent Server (http:// patent.womplex.ibm.com) offers several advantages over the USPTO patent site from which it derives its information, particularly ease-of-use and free full-patent images. The site directly links to Optipat, which sells full-text patents for as little as $2.50 each.

MicroPatent and its Internet-based patent site (http://www.micropat.com) offer other advantages. MicroPatent takes a step forward by offering a fee-based e-mail alerting service based on patent applications filed with the European Patent Office (EPO), World Intellectual Property Organization (WIPO), and Japanese Patent Office (JPO). This service can offer a competitor intelligence advantage: U.S. patent information is only released when a patent is granted -- often several years after the filing of the application. EPO, WIPO, and JPO, on the other hand, release information within 18 months of the application filing -- sometimes providing a competitive leg up, so to speak.

Small Is Beautiful

While we often expect major new services to provide the greatest utility, sometimes the small things offer the greatest value. For many of our research projects at Fuld, we routinely retrieve SEC filings, seeking invaluable information concerning financials, operations, products sold, markets, management, and much more. Competitors' SEC filings are fundamental competitor intelligence

research tools.

Several companies now offer inexpensive or free SEC- and Internet-based alerting services. Who Where? (http://www.whowhere. com) sponsors EDGARAlert! (http://www.whowhere. com/EDGAR), a free service that sends you e-mail when one of your specified companies submits a new SEC filing, though there may be a time delay. In contrast, EDGAR Online's Watchlist (http://www.edgar-online.com/) service is a low-cost, real-time alerting service. As soon as the SEC filing is submitted, you receive an HTML-enabled alert in your mailbox. Click on the hyperlink and you go directly to the filing in question. It also features a new service called EDGAR Online People that allows customers to search SEC filings by an executive's name and thus uncover position held within the corporation, compensation packages, and board memberships.

The Help-Wanted Intelligence Explosion

While numerous Internet services abound offering inexpensive alternatives to the "high-priced spreads" of commercial online services, the Internet does offer new and unique services. For example, job postings provide a wealth of clues to a company's current technology use, research interests, and future plans. Identify the persons and skills your competitors hire today and you will often uncover their future. In years past, this information was only available through specialty database companies, newspaper clipping services, or by subscribing to newspapers in competitors' home towns. Today, a myriad of ways can help identify hiring trends, including competitors' home pages and searching Usenet group posting via Deja News (http://www.dejanews.com).

Commercial job sites such as Monsterboard (http://www.monsterboard. com) and CareerMosaic (http://www. careermosaic.com) can be worth their weight in gold. Companies frequently post lengthy lists of various and varying jobs, offering a peek at the range of their work. Monsterboard now offers an alerting service aimed at job seekers that can also help the CI researcher. You can specify company name or type of job and receive an e-mail alert when that company posts a job.

For smaller companies, or those not yet Internet-savvy, CareerPath (http:// career.careerpath.com) is a unique resource which expanded coverage last year. Several years ago, CareerPath pioneered Internet job postings by offering two weeks of job advertisements from six major newspapers across the country, searchable via one interface. The site still allows a two-week search by company name or job category, but now includes more than 40 large and small newspapers across the country and imports job postings from major corporate Web sites.

Many people err in believing that Japan is our largest trading partner, when in fact that honor belongs to Canada. Thus, a welcome addition to business research on the Internet was last

year's start-up of Canada's corporate filing site, similar to the U.S. EDGAR site. Called SEDAR, (System for Electronic Document Analysis and Retrieval), the bilingual service is free and includes most filings required of Canadian public companies going back to January 1, 1997 (http://www.sedar.com/).

Last year also saw the death of another valuable commercial service known for unusual and sometimes unique information. NewsNet, noted for its collection of hard-to-find industry newsletters, including some newsletters unique to its service, ceased operation in August 1997. Its remaining Web site now prompts former subscribers to a Web-based online service called BrainWave, offered by WinStar Telebase. This service features an interface that mimics NewsNet's old Baton interface. However, BrainWave's newsletter offerings are limited to those included in IAC databases. Due to NewsNet's demise, several technical newsletters are no longer available electronically -- commercially or via the Internet -- a tremendous loss for researchers in those industries.

And Now 1998...

As to the future for competitor intelligence and online resources? Remember that line from The Graduate: "Plastics"? Just do a global replace with the word "Internet." A clear message has been sent: Proprietary systems will see few new developments. Look at the highly visible migration to the Web with DialogWeb and Dow Jones Interactive. The Internet is the clear winner of the platform race.

When used to its full potential, the Internet offers competitive advantages over older systems. Expect traditional market forces to apply, however. New entrants must offer clear advantages. In the Internet world, creativity is far easier to accomplish when you are a brash upstart, not hamstrung by commitments to older technologies. Look to the rebels for sizzling innovative products offering such desirable features as intuitive interfaces, access to new resources, unique services, time savings, lower costs, or a combination of the above. We will all adjust to the thrills, challenges, and exasperation inherent in this brave new Internet world.
